

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 02:49:54 ; Search time 4459 Seconds
(without alignments)
12395.003 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaaagcaacgt.....gaatttcacaaacattga 1452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1452	100.0	1602	3	CNS0A0KS	BX833171 Arabidops
2	1450	99.9	1656	3	CNS0A1A4	BX829836 Arabidops
3	1446.8	99.6	1585	3	CNS0A1A4	BX829783 Arabidops
4	606.8	41.8	624	1	AV823628	AV823628 AV823628
5	599.6	41.3	606	6	CB261053	CB261053 35-E9570-
6	406.4	28.0	920	7	CF215873	CF215873 CAST0002
7	391	26.9	894	7	CO117860	CO117860 GR_Eb01O
8	371	25.6	866	7	CO095556	CO095556 GR_Ea18E
9	364.6	25.1	749	7	CN187799	CN187799 UCRC05_0
10	357.2	24.6	723	7	CK319497	CK319497 X9P11dhl
11	355.2	24.5	825	7	CO096428	CO096428 GR_Ea19K
12	347	23.9	348	7	Z37245	Z37245 ATTS3988 St
13	343.2	23.6	2173	3	AY596550	AY596550 Saccharum
14	339.8	23.4	742	7	CV197356	CV197356 CGF100395
15	338.4	23.3	2189	3	AY105375	AY105375 Zea mays
16	330.2	22.7	977	7	CK283827	CK283827 EST746549
17	328.6	22.6	886	7	CK289726	CK289726 EST752448
18	325.8	22.4	849	7	CK264004	CK264004 EST710082
19	324	22.3	883	7	CO073048	CO073048 GR_Ea32I
20	323.8	22.3	681	5	BUB36279	BUB36279 T08402 P
21	315.2	21.7	608	7	CN905453	CN905453 010928ABD
22	313.6	21.6	613	7	CN913815	CN913815 030108ABM
23	313.6	21.6	636	7	CN913251	CN913251 030107ABM
24	312	21.5	663	1	AJ805795	AJ805795 AJ805795

25	311.8	21.5	763	2	BE034121	BE034121 MG05C04 M
26	307.2	21.2	594	7	CN912222	CN912222 021217ABM
27	307.2	21.2	626	7	CN862200	CN862200 000818AAL
28	305.2	21.0	721	6	CB917703	CB917703 VVD019F01
29	301.2	20.7	586	1	AI998258	AI998258 701544952
30	298.2	20.5	688	7	CN884491	CN884491 010811AAS
31	294.8	20.3	717	7	CF808960	CF808960 PSHB041XE
32	294.8	20.3	764	5	BQ045892	BQ045892 EST595010
33	293	20.2	496	1	AV784632	AV784632 AV784632
34	292.8	20.2	796	7	CK261683	CK261683 EST707761
35	290	20.0	290	1	AV525150	AV525150 AV525150
36	283.8	19.5	536	5	BQ588266	BQ588266 S013218W-
37	281.8	19.4	559	4	BI471389	BI471389 sag19h04.
38	281.2	19.4	497	7	CN865160	CN865160 001001AAL
39	280.8	19.3	500	2	AW201710	AW201710 sf06d03.Y
40	280.4	19.3	600	4	BG139883	BG139883 EST480325
41	280.2	19.3	504	1	AI988515	AI988515 sd03d02.Y
42	280.2	19.3	518	4	BG882882	BG882882 sae78h02.
43	280.2	19.3	542	5	BQ080012	BQ080012 san04f04.
44	280.2	19.3	849	7	CV287071	CV287071 68153.1 A
45	279.2	19.2	583	2	AW686368	AW686368 NF040G09N

ALIGNMENTS

RESULT 1
CNS0A0KS 1602 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from Clone GSITSL37ZE02 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION BX833171.1 GI:42457670
VERSION HTC; GSLT cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1602)
AUTHORS Castellani V., Aury J.M., Jaillon O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpetti C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1602)
AUTHORS Genoscope.
JOURNAL Direct Submission
TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Sequenage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr) - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/external/sequences/Banque_Projet_EF/Full length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source
Location/Qualifiers
1..1602
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

```
/clone="GSLTSL372E02"
/tissue_type="Silique"
/plasmid="pCWSFOR6"
/complement(i..1602)
/gene="At5g54310"

ORIGIN
Query Match 100.0%; Score 1452; DB 3; Length 1602;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCGCGCATAGAAAGATCTTGA 60
Db |||||||
Qy 75 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCGCGCATAGAAAGATCTTGA 134
Db |||||||
Qy 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCACAGA 120
Db |||||||
Qy 135 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCACAGA 194
Db |||||||
Qy 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
Db |||||||
Qy 195 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 254
Db |||||||
Qy 181 CTGGGGGTACACATATCGAAGGTTCCGATCTGCACTCTGGACACATGGCTCCCGAGCAG 240
Db |||||||
Qy 255 CTGGGGGTACACATATCGAAGGTTCCGATCTGCACTCTGGACACATGGCTCCCGAGCAG 314
Db |||||||
Qy 241 GTTGCAATTATACAGTCAATGGAATGATAAGCAAAATAGTTACTTGGGAAGCGAGCTA 300
Db |||||||
Qy 315 GTTGCAATTATACAGTCAATGGAATGATAAGCAAAATAGTTACTTGGGAAGCGAGCTA 374
Db |||||||
Qy 301 CCCCCAAACTATGATAGAGTTGGATTCGAGATTTTATACGTGCAAGATATGAGAGAG 360
Db |||||||
Qy 375 CCCCCAAACTATGATAGAGTTGGATTCGAGATTTTATACGTGCAAGATATGAGAGAG 434
Db |||||||
Qy 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGACAGCAAGCGGG 420
Db |||||||
Qy 435 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGACAGCAAGCGGG 494
Db |||||||
Qy 421 AAATCTGTGGAGAGAGTGGGCGGGATGATGAGCATGAGCATATAGTAGTCTGTAAAT 480
Db |||||||
Qy 495 AAATCTGTGGAGAGAGTGGGCGGGATGATGAGCATGAGCATATAGTAGTCTGTAAAT 554
Db |||||||
Qy 481 TTGTTTGGAGGAGAGAAACTATTCAGCATCTAGCAAGAAATAAATGTTGCTGCAAG 540
Db |||||||
Qy 555 TTGTTTGGAGGAGAGAAACTATTCAGCATCTAGCAAGAAATAAATGTTGCTGCAAG 614
Db |||||||
Qy 541 AGAATAAATCTTCCCGTGTCTCCCAAGGACCCAGTCTAGGTTATAAGCCACAGCAAAA 600
Db |||||||
Qy 615 AGAATAAATCTTCCCGTGTCTCCCAAGGACCCAGTCTAGGTTATAAGCCACAGCAAAA 674
Db |||||||
Qy 601 ATGGAGTCTGACGATCTCCAGTAGAGAGGAGAAACAGCGATTAATGTTGGACACGCA 660
Db |||||||
Qy 675 ATGGAGTCTGACGATCTCCAGTAGAGAGGAGAAACAGCGATTAATGTTGGACACGCA 734
Db |||||||
Qy 661 TCAGATCTCCAAAGGTGATTTCTGATCTGATCTGTTTAAACATGCTCAATGATGAT 720
Db |||||||
Qy 735 TCAGATCTCCAAAGGTGATTTCTGATCTGATCTGTTTAAACATGCTCAATGATGAT 794
Db |||||||
Qy 721 TCGACTACAAATACCTCAGAGGCAACTCTGCGGATACCTCTGCGGATGATAACTCATGG 780
Db |||||||
Qy 795 TCGACTACAAATACCTCAGAGGCAACTCTGCGGATACCTCTGCGGATGATAACTCATGG 854
Db |||||||
Qy 781 GCTGCTTTTCACTGTGCGAAGTGGTCAACCGGAGAGAAATTTGTCAAGCCAGCCT 840
Db |||||||
Qy 855 GCTGCTTTTCACTGTGCGAAGTGGTCAACCGGAGAGAAATTTGTCAAGCCAGCCT 914
Db |||||||
Qy 841 GCTGAGAGAGTTCCTCCAGCTTCATCTTCTGATCTTGGATTTCTTTAAGACACA 900
Db |||||||
Qy 915 GCTGAGAGAGTTCCTCCAGCTTCATCTTCTGATCTTGGATTTCTTTAAGACACA 974
Db |||||||
Qy 901 CCTAATTTAACTCAACAGCACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
Db |||||||
```

```
Db 975 CCTAATTTAACTCAACAGCACAAAGATGTGAAGCGGATATCATGAGCCTGTTT 1034
Qy 961 GAGAAGAGCAATATAGTATGCGCTTTTGGCCATGATCAGCAACAGGTTGCTATGCTCGCT 1020
Db |||||||
Db 1035 GAGAAGAGCAATATAGTATGCGCTTTTGGCCATGATCAGCAACAGGTTGCTATGCTCGCT 1094
Qy 1021 CAGCAGCAAGCCCTTTTACATGGCTGCGAGCGAAAGCTGCTGGAGGCACTCCAAAAGCGCGTG 1080
Db |||||||
Db 1095 CAGCAGCAAGCCCTTTTACATGGCTGCGAGCGAAAGCTGCTGGAGGCACTCCAAAAGCGCGTG 1154
Qy 1081 AATCAACAAAGCTATTTGCTTAATGCTCTTTAAAGTCTTCTGCAAAATGCTCAAAAGCGCGC 1140
Db |||||||
Db 1155 AATCAACAAAGCTATTTGCTTAATGCTCTTTAAAGTCTTCTGCAAAATGCTCAAAAGCGCGC 1214
Qy 1141 GGCTACAGATCCCGGAATGACTTAACCCGTTAGGTGCTCAAGCTGATCTCCAGAACTT 1200
Db |||||||
Db 1215 GGCTACAGATCCCGGAATGACTTAACCCGTTAGGTGCTCAAGCTGATCTCCAGAACTT 1274
Qy 1201 ATGCAAAACATGAATATGAACGCAAAACATGAACACAGAGACCCGCAACCCGCAAGAGAAC 1260
Db |||||||
Db 1275 ATGCAAAACATGAATATGAACGCAAAACATGAACACAGAGACCCGCAACCCGCAAGAGAAC 1334
Qy 1261 ACTCAAAATACCCATCATCCAGTTTCTTCAACAATGGGTCAAGTCAATCAAGTGAACGGT 1320
Db |||||||
Db 1335 ACTCAAAATACCCATCATCCAGTTTCTTCAACAATGGGTCAAGTCAATCAAGTGAACGGT 1394
Qy 1321 ATGACCCCAACCTCAACCGGTAAACCTCAGTCAATCCGCAACCCCAACCAACCAAGCACC 1380
Db |||||||
Db 1395 ATGACCCCAACCTCAACCGGTAAACCTCAGTCAATCCGCAACCCCAACCAACCAAGCACC 1454
Qy 1381 ACACCATCTTCACAATCAGGCAAAAGACTTTGATTTCTTCTTCTTGTGATGGAATGTTTC 1440
Db |||||||
Db 1455 ACACCATCTTCACAATCAGGCAAAAGACTTTGATTTCTTCTTCTTGTGATGGAATGTTTC 1514
Qy 1441 ACAAAACATTGA 1452
Db 1515 ACAAAACATTGA 1526
```

```
RESULT 2
CNS0A1A4 1656 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GILT32H12 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX829836
VERSION BX829836.1 GI:42458516
KEYWORDS HTC; GSLT.cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1656)
AUTHORS Castellani, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
TITLE Whole Genome Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1656)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellani
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
```


JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 1585)
Genoscope.

Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Pull
length

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

Location/Qualifiers
1. .1585
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLFB36ZH03"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT 6"
complement(1..1585)
/gene="At5g54310"

FEATURES

source

gene

ORIGIN

Query Match 99.8%; Score 1446.8; DB 3; Length 1585;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1448; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	3	GAACGAGAAAGCCAAAGCTCTTAAGAGAGCTTAATGCCGCCCATAGAGAAATTTCTTGAAGG	62
Db	1	GAACGAGAAAGCCAAAGCTCTTAAGAGAGCTTAATGCCGCCCATAGAGAAATTTCTTGAAGG	60
QY	63	GCTTCTTAAACATCCAGAGAACAGAGATGTGCTGATCGCAAAACAAAGGTCCAAAGATG	122
Db	61	GCTTCTTAAACATCCAGAGAACAGAGATGTGCTGATCGCAAAACAAAGGTCCAAAGATG	120
QY	123	GGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGATCT	182
Db	121	GGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGATCT	180
QY	183	CGGGGTACACATATCGAAGGTTGATCTGCTGCTGACATCTGGACACATGGCTCCCGAGCAGGT	242
Db	181	CGGGGTACACATATCGAAGGTTGATCTGCTGCTGACATCTGGACACATGGCTCCCGAGCAGGT	240
QY	243	TGCATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGTACC	302
Db	241	TGCATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGTACC	300
QY	303	CCCAACTATGATAGAGTTGGAAATTCAGAAATTTATACGTGCAGAGTATGAGAGAGAG	362
Db	301	CCCAACTATGATAGAGTTGGAAATTCAGAAATTTATACGTGCAGAGTATGAGAGAGAG	360
QY	363	ATGGGTTTCTAGAGGGGAAAGGCTAGATCCTCTCTAGAGTCGACGAGGAACGGCGGAA	422
Db	361	ATGGGTTTCTAGAGGGGAAAGGCTAGATCCTCTCTAGAGTCGACGAGGAACGGCGGAA	420
QY	423	ATCTGTGGAGAGAGTGGGCGGGATATGAGCATGACATAGTAGTGTCTGTAAATTT	482
Db	421	ATCTGTGGAGAGAGTGGGCGGGATATGAGCATGACATAGTAGTGTCTGTAAATTT	480
QY	483	GTTTGGAGAGAGAAACTATTTCCAGCATCTAGAACAGAAATATGTTGCTGCAACGAG	542
Db	481	GTTTGGAGAGAGAAACTATTTCCAGCATCTAGAACAGAAATATGTTGCTGCAACGAG	540

QY	543	AATAAATCTTCCCGTCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAAAAT	602
Db	541	AATAAATCTTCCCGTCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAAAAT	600
QY	603	GGAGTCTGACGTACTCCAGTAGAGAGGGAGAAAACAAGCAGTAATGTTGACACGATC	662
Db	601	GGAGTCTGACGTACTCCAGTAGAGAGGGAGAAAACAAGCAGTAATGTTGACACGATC	660
QY	663	AGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAAACATGCTTATCAATGGATGATTC	722
Db	661	AGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAAACATGCTTATCAATGGATGATTC	720
QY	723	GACTACAAATACCTCAGAGGCAACTCTCGGCGATCTCTCCGATGATACTCATGGC	782
Db	721	GACTACAAATACCTCAGAGGCAACTCTCTGGCGATCTCTCTCCGATGATACTCATGGC	780
QY	783	TGGCTTTTCAGTCTCTGGAAGTGTCAAACGGCAGAGAAATTTCTCAGACCCACCTGC	842
Db	781	TGGCTTTTCAGTCTCTGGAAGTGTCAAACGGCAGAGAAATTTCTCAGACCCACCTGC	840
QY	843	TGAGAGCAGTCTCTCTCCAGCTTCATCTTCTGACTTTTGAGGATTTGTTTAAAGGACACACC	902
Db	841	TGAGAGCAGTCTCTCTCCAGCTTCATCTTCTGACTTTTGAGGATTTGTTTAAAGGACACACC	900
QY	903	TAATTTAACTCAACAGCACCAAAAGATGTGAAAGGCGATATCATGAGCCTGTTGA	962
Db	901	TAATTTAACTCAACAGCACCAAAAGATGTGAAAGGCGATATCATGAGCCTGTTGA	960
QY	963	GAAAGCAGATATGATATGCGCTTTTGCCATGATCAGCAACAGTTGCTGCTCGCTCA	1022
Db	961	GAAAGCAGATATGATATGCGCTTTTGCCATGATCAGCAACAGTTGCTGCTCGCTCA	1020
QY	1023	GCAGCAGCCCTTTTACATGGCTGCAGGAAAGCTGCTGGAGGCACTCCAAACGGCGTGA	1082
Db	1021	GCAGCAGCCCTTTTACATGGCTGCAGGAAAGCTGCTGGAGGCACTCCAAACGGCGTGA	1080
QY	1083	TCAAAGCTATTTGCTAATGCTCTTAAAGCTAGCTTTGCAAAATTTGGTCAAACCCCGCGG	1142
Db	1081	TCAAAGCTATTTGCTAATGCTCTTAAAGCTAGCTTTGCAAAATTTGGTCAAACCCCGCGG	1140
QY	1143	CTACAGATCCCGGAAATGACTTACCCGTAGGTGCTCAAGCTGATCTCCGAAACTTAT	1202
Db	1141	CTACAGATCCCGGAAATGACTTACCCGTAGGTGCTCAAGCTGATCTCCGAAACTTAT	1200
QY	1203	GCAAAACATGATATGAACGCAAAATGAAACGAGACCCGCAACACCGCAAGAGAACAC	1262
Db	1201	GCAAAACATGATATGAACGCAAAATGAAACGAGACCCGCAACACCGCAAGAGAACAC	1260
QY	1263	TCTACAATACCCATCATCCAGTTTCTACAATGGGTCAAGCTAATCAAGTGAACGGTAT	1322
Db	1261	TCTACAATACCCATCATCCAGTTTCTACAATGGGTCAAGCTAATCAAGTGAACGGTAT	1320
QY	1323	GACCCCAACTCAACCGGTAAACCTCAGTCAATATCGCAACCCCAACCAAGCACAC	1382
Db	1321	GACCCCAACTCAACCGGTAAACCTCAGTCAATATCGCAACCCCAACCAAGCACAC	1380
QY	1383	ACCATCTTCAATCAGGCAAGACTTTGATTTCTTCTTGTGATGGAATGTTTCAAC	1442
Db	1381	ACCATCTTCAATCAGGCAAGACTTTGATTTCTTCTTGTGATGGAATGTTTCAAC	1440
QY	1443	AAAACATTGA 1452	
Db	1441	AAAACATTGA 1450	

RESULT 4
AV823628

LOCUS

DEFINITION

AV823628 RAPUS Arabidopsis thaliana cDNA clone RAPL05-21-P04 5',

624 bp mRNA linear EST 01-APR-2002

AV823628 mRNA sequence.

AV823628

VERSION

AV823628.1 GI:19865688

KEYWORDS

EST.

SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 624)

REFERENCE
AUTHORS
Seki, M., Narusaka, M., Iehida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified paluascript vector as a Set1/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
further details.

FEATURES
source
Location/Qualifiers
1..624
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-21-P04"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/notes="Site 1: Set1; Site 2: XhoI; subjected to
dehydration-treated (1,2,5,10,24 hr)"

ORIGIN
Query Match 41.8%; Score 606.8; DB 1; Length 624;
Best Local Similarity 99.3%; Pred. No. 1.7e-166;
Matches 608; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAACCAACGCTCTTAAGAGAGCTTAATGCCCGCCATAGAAAGATTTCTTGA 60
DB 13 ATGACGAGAGAGCAACGCTCTTAAGAGAGCTTAATGCCCGCCATAGAAAGATTTCTTGA 72
QY 61 GGCTTTCTTAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTCCAG 120
DB 73 GGCTTTCTTAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTCCAG 132
QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCATGCTGCTGGATTCACAGAGT 180
DB 133 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCATGCTGCTGGATTCACAGAGT 192
QY 181 CTCGGGTACACATATCGAAGTTGATCTGCCACTCTGGACATCGCTCCCGAGCAG 240
DB 193 CTCGGGTACACATATCGAAGTTGATCTGCCACTCTGGACATCGCTCCCGAGCAG 252
QY 241 GTTGCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCT 300
DB 253 GTTGCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCT 312
QY 301 CCCCAAACTATCATAGAGTTGGAATTTGCAATTTTATAGTCCCAAGTATGAAGAGAG 360
DB 313 CCCCAAACTATCATAGAGTTGGAATTTGCAATTTTATAGTCCCAAGTATGAAGAGAG 372
QY 361 AGATGGTTTCTAGAGGGGAAAGGCTAGATCACTCTAGAGTCGAGGAGGAGCGCG 420
DB 373 AGATGGTTTCTAGAGGGGAAAGGCTAGATCACTCTAGAGTCGAGGAGGAGCGCG 432
QY 421 AAATCTGTGAGAGAGAGTGGCGCGGATATGAGCATGGACATAGTAGTCTCTGTAAT 480
DB 433 AAATCTGTGAGAGAGAGTGGCGCGGATATGAGCATGGACATAGTAGTCTCTGTAAT 492

QY 481 TTGTTTGAGGAGAGAAACTATTCCAGCATCTAGACAAAGAAATAATTTCTGCAACG 540
DB 493 TTGTTTGAGGAGAGAAACTATTTCAGCATCTAGACAAAGAAATAATTTCTGCAACG 552
QY 541 AGAATAAATCTTCCGTCCTCCCAAGACCCAGTTCAGTTTATAAGCCAGCAGAGAA 600
DB 553 AGAATAAATCTTCCGTCCTCCCAAGACCCAGTTCAGTTTATAAGCCAGCAGAGAA 612
QY 601 ATGGAGTCTGCA 612
DB 613 ATGGAGTCTGCA 624

RESULT 5
CB261053
LOCUS
DEFINITION
35-E9570-012-004-E09-T7R MPZ-ADIS-012 Arabidopsis thaliana cDNA
clone MPZp769E094Q 5-PRIME, mRNA sequence.
CB261053
ACCESSION
CB261053.1 GI:32885826
VERSION
EST.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 606)
Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Altmann, T.,
Mitchell-Olds, T. and Weisshaar, B.
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
Genome Res. 13 (6), 1250-1257 (2003)
22683290
12799357
Contact: Weisshaar B
ADIS DNA core facility at MPZ
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@mpiz-koeln.mpg.de
Insert Length: 606 Std Error: 0.00
Plate: 4 row: E column: 09
Seq primer: T7R; CTAATCGACTCATATAGGA.
Location/Qualifiers
1..606
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Landsberg erecta"
/db_xref="GABI:589856"
/db_xref="taxon:3702"
/clone="MPZp769E094Q"
/tissue_type="whole plant"
/dev_stage="adult plant, mixed stresses"
/lab_host="E. coli XL1-Blue MRF"
/clone_lib="MPZ-ADIS-012"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
library from Arabidopsis thaliana, accession Landsberg
erecta; six weeks old total plants grown under long-day
conditions in soil, whole adult plants were treated for 24
hours with different stresses, (1) at 4M-0 C in the dark,
(2), at 37 Grad C in the dark, (3) lying in the lab after
removing from soil, (4) in the greenhouse after wounding
leaves with a forceps, (5) in the lab watering with a 150
mM NaCl solution, (6) at 26 Grad C in the light/UV; equal
quantities of stressed plant material were pooled; library
was made at the Max-Planck-Institute for Plant Breeding
Research, Cologne, Germany; cloning sites SalI-NotI,
primer sites and orientation:
T7-SalI-CCACGCTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
Sequencing granted in the context of the GABI Arabidopsis
Verbund I: Genetic Diversity, 'Establishment of
high-efficiency SNP-based mapping tools and development of
methods for genome-wide mutation detection' PI: Bernd

COMMENT

Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 18 row: B column: 01.
Location/Qualifiers
1. .866
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eal8E01"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="Vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional clones into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 25.6%; Score 371; DB 7; Length 866;
Best Local Similarity 73.1%; Pred. No. 3.8e-97;
Matches 505; Conservative 0; Mismatches 180; Indels 6; Gaps 2;

QY 1 ATGACGAGAGGCAACGCTCTTAAGGAGCTTAATGCCGCCATAGAAAGATCTTGAA 60
DB |||||
QY 179 ATGAACGAGAGGCGAGCGTTACCAAGGAGCTCAACGCCAGACCCGAAGATTATTGAA 238
DB |||||
QY 61 GGGCTCTTTAAACATCCAGAGAACAGAGAATGTGCTGACTGCAAAACAAAGAGTCCAAGA 120
DB |||||
QY 239 AGTCTCTTAATTCGCGGAGATAGGAGATGGCGGACTGCAAGCCAAAGGTCGCGA 298
DB |||||
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGGAGT 180
DB |||||
QY 299 TGGGCAAGTGTGAATTTAGGTATCTTTATGATGATGCAATGTTCTGGGATTCACAGGAGT 358
DB |||||
QY 181 CTGGGGTACACATATCAAGGTTGATGCTCCACTCTGACACATGCTCCCGGACG 240
DB |||||
QY 359 CTGGGGTACACATATCAAGGTTGATGCTCTGCTACATAGACATGCTCTTCTGGAGCAG 418
DB |||||
QY 241 GTTGCAATTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB |||||
QY 419 GTTGCTTTTATTCATCAATCGGGATGAAAGGCAACAGTACTGGAGCTGAGTTA 478
DB |||||
QY 301 CCCCCAAATATGATAGAGTTGGAAATGAGAAATTTTATCTGTCGAAAGTATGAAGAGAAG 360
DB |||||
QY 479 CCCCCCTAATATGATAGAGTTGGAAATGAGAAATTTTATCTGTCGAAAGTATGAAGAGAAG 538
DB |||||
QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTAGATGTCGACGAGGACGCGGG 420
DB |||||
QY 539 AGATGGGTTCTAGAGAGATGGAATTTCCAAATCACCACCTGAGAGGTTGGACGAAAGGCT 598
DB |||||
QY 421 AAATC---TGTGGAGAGAGTGGGCGGGATATGAGCATGATGACATAGTAGTAGTCTGTA 477
DB |||||
QY 599 CCCTCAGATGTCGAGACCTAATGAAACAGTGTCTATGGGCACATAGTAATTCGAG 658
DB |||||
QY 478 AATTGTTTGGAGGAGGAAACTATTTCAGCATCTAGAACAGAAATAATGTTGCTGCA 537
DB |||||
QY 659 AATTCATTTGAGGAAGGAGGACAAACAAAGCACTTGTCTAGAAAGAAATCTTCTGCA 718
DB |||||
QY 538 ACGAGAAATAATTTCCCGTGCCTCCCAAGACCCAGTCAAGTTATTAAGCCACAGCAG 597
DB |||||
QY 719 ACAGAGTTAGTCTTCTGTTCTCTCTTCAAGGACCTGATCAGGTTACA---CCGCTACAG 775
DB |||||
QY 598 AAATGAGTCTGACGACTCTCCAGTAGAGGGAGAACAGCAGTAAGTTGTCACCA 657
DB |||||
QY 776 AAGCCTGAACCAAGTTGTTGCGCCAGCTGAGGCAACAAAGCCAGCTGTAGAGACTGCTCCC 835
DB |||||
QY 658 GCATCAGATCTCCAAAGGTTGATTTTCTGA 688

Db 836 GTAGCCATCGCTCTCTAAAGTCGATTATGCTA 866

RESULT 9
CN187799
LOCUS
DEFINITION
Library Citrus sinensis cDNA clone CS_WEB0005A10, mRNA sequence.
CN187799
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 749)
AUTHORS
Close, T.J., Rose, M.L., Arpaia, M.L., Federici, C.F., Fenton, R.D.,
Wanamaker, S., Focht, E., Sievert, J., Robinson, P., Kim, H.R.,
Kudrna, D., Stum, D., Yost, D. and Wing, R.
TITLE
Development of EST Resources and New Genetic Markers for California
Citrus - Washington Navel Orange Stored Fruit Pulp
JOURNAL
Unpublished (2004)
COMMENT
Contact: Timothy Close
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124
Tel: 9097873318
Fax: 9097874437
Email: timothy.close@ucr.edu
Seq primer: T3.
Location/Qualifiers
1. .749
/organism="Citrus sinensis"
/mol_type="mRNA"
/cultivar="Washington navel"
/db_xref="taxon:2711"
/clone="CS_WEB0005A10"
/tissue_type="Pulp"
/dev_stage="Commercially producing trees"
/lab_host="E. coli TJCL21"
/clone_lib="Washington Navel Orange Stored Fruit Pulp cDNA
Library"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Fruits were collected
January-March 2003 (Federici, Rose lab; Focht, Sievert &
Robinson, Arpaia lab). Four samples related to storage
conditions were produced: 1) fresh-picked in Mentone
(Arnott Brothers Enterprises, Mentone, CA), 2) after 21
days storage at 5C at Kearney then transported to UC
Riverside on ice, 3) after 5 additional days storage at
11C at Kearney, sampled immediately, 4) fruit grown in
southern CA were obtained from Redlands Foothill Packing
House after commercial packing, X-ray irradiated at 300 Gy
by Surebeam, then stored 1 day at ambient temperature.
Pulp tissue (juice vesicles) were collected. Tissues were
snap frozen and then stored at -80C until further
processing. Fenton (Close lab) purified RNA by the phenol
method described in J. Japanese Soc. Hort. Sci. 1996. 64
(4): 809-814, purified poly(A) mRNA using a PolyAtrack
mRNA Isolation System IV (Promega), produced a primary
cDNA library using a lambda ZAP XR cDNA Synthesis Kit
(Stratagene), then mass-excised one million pfu from the
primary library to produce a phagemid population.
Phagemids were plated, plasmid DNA purified, cDNA clones
archived, and DNA sequences determined bi-directionally
using an ABI3730 at the Arizona Genomics Institute,
University of Arizona (Kim, Kudrna, Stum, Yost, Wing).
Chromatogram files were downloaded by FTP to UC Riverside
(by Close), then processed at UC Riverside (by Wanamaker,
Close lab) using the HarVest pipeline
(<http://harvest.ucr.edu>) to remove vector and cloning
oligo sequences and various contaminants, and to trim to

FEATURES
source

high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."

```

ORIGIN
Query Match      25.1%; Score 364.6; DB 7; Length 749;
Best Local Similarity 77.5%; Pred. No. 2.7e-95;
Matches 455; Conservative 0; Mismatches 129; Indels 3; Gaps 1;

QY 1 ATGAACGAGAACCCACGCTCTTAGAGAGCTTAATGCCCGCATAGAAAGATTTCTTGA 60
DB 104 ATGAACGAGAGGCCAACGCTCTTAAGAGCTTAACGCTAGACACAGGAAGATTTCTGAA 163
QY 61 GGCTCTTTAAACATCCAGAGACAGAGATGTGCTGACTGCAAAAACAAAAGTCCAAAG 120
DB 164 GGCTCTTTAAAGTTCAGAGATAGAGGATGGCTGACTGCAAGCCAAAGTCCAAAG 223
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATCAATGTTCTGGGATTCACAGAGT 180
DB 224 TGGGCTAGTGTGAATTTGGGAATCTTTATCTGCATCAATGTTCTGGGATTCACAGAGT 283
QY 181 CTGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
DB 284 CTGGGTATACACATATCAAGGTTAGATCTGTACCTTGGATACCTGGCTCCGAGCAG 343
QY 241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB 344 GTTGCAATTTCAATCTATGGGAATAGAGAGCAAAATAGTTACTGGGAAGTCTGAGTTG 403
QY 301 CCCCCAACTATCATAGAGTTGGAATTTGAGAAATTTTATAGTGCACAAAGTATGAAGAGAG 360
DB 404 CTTCCAAACTATCAGAGTTGGAATTTGAGAAATTTTCAATCTGTCGCAAGTACGAGGAGAG 463
QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAG---TCGACAGGAGACGG 417
DB 464 CGATGGGTTTCTAGAGATGACAAAGCAAAATCACTCTAGAGGGCTGGAGGAGAGGCT 523
QY 418 CGGAATCTGTGAGAGAGAGTGGCGGGATATGAGCATGGACATAGTAGTACTCTCTGTA 477
DB 524 TCTATTCTATGGCAGAGACTGGGGAAGAGAGTGGGATGGTTACTGANTATTTCTGAA 583
QY 478 AATTTGTTTGGAGAGAGGAAACTATTCCAGCATCTAGAACAAAGAAATATGTTGCTGCA 537
DB 584 AACTTATCTGAGGAAGGAGGATGTTCAAGCTCCAAGTACAAAGATAGTGTCTCTGCT 643
QY 538 ACAGAAATTAATCTCCGCTGCTCCCAAGGACCCAGTCAGGTTAT 584
DB 644 GCAAGAAATGGTCTTCCACTTCCACTTAGAGGACCTGATCAGGTTGT 690

RESULT 10
CK319497
LOCUS CK319497.1 GI:47107920
DEFINITION X911d11 Populus stem seasonal library Populus deltoides cDNA, mRNA
sequence.
ACCESSION CK319497
VERSION CK319497.1
KEYWORDS EST.
SOURCE Populus deltoides
ORGANISM Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
1 (bases 1 to 723)
Park, S. and Han, K.-H.
Gene expression profile during seasonal growth cycle in poplar tree
Unpublished (2003)
Contact: Kyung-Hwan Han
Department of Forestry
Michigan State University
126 Natural Resources, East Lansing, MI 48824-1222, USA
Tel: 517 353 4751
Fax: 517 432 1143
Email: hanky@msu.edu.

REFERENCE
AUTHORS CK319497
TITLE X911d11 Populus stem seasonal library Populus deltoides cDNA, mRNA
JOURNAL
COMMENT
```

FEATURES

```

source
1..723
Location/Qualifiers
/organism="Populus deltoides"
/mol_type="mRNA"
/strain="ILL-129"
/db_xref="taxon:3696"
/tissue_type="stem"
/dev_stage="1 year old"
/clone_lib="Populus stem seasonal library"

ORIGIN
Query Match      24.6%; Score 357.2; DB 7; Length 723;
Best Local Similarity 71.0%; Pred. No. 4e-93;
Matches 508; Conservative 0; Mismatches 189; Indels 18; Gaps 2;

QY 19 GTCTCTAAGGAGCTTAATGTCGCCCATAGAAAGATTTCTTGAAGGGCTCTTAAACATCA 78
DB 4 GTCTCTAAGAGCTCAATGCTAGACACAGAAAGATTTCTTGAAGGGCTCTCAAAATGGCT 63
QY 79 GAGAACAGAGAAATGTGCTGACTGCAAAAACAAAAGTCCAAAGTGGGCTAGTGTAAATTA 138
DB 64 GAGATAGGAAATGTGCTGACTGCAAAAGTCCAAAGTGGCAAGCGTGAATTTG 123
QY 139 GGTATCTTTATCTGCATCAATGTTCTGGATTCACAGAGTCTCGGGGTACACATATCG 198
DB 124 GGTATCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 183
QY 199 AAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAGGTTGCATTTATACAGTCA 258
DB 184 AAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAGGTTGCATTTATTCATCA 243
QY 259 ATGGGAAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATGA 318
DB 244 ATGGGAAATGAGAGGCAAAATAGTTATTTGGGAAGCAGAGTTACCCCAAACTACGACA 303
QY 319 GTTGAATTTGAAATTTTATACGTGCAAAAGTATGAAGAGAGAGATGGGTTCTTAGAGG 378
DB 304 GTTGAATTTGAAATTTTCAATCTGCGAAGTATGAAGAGAGAGATGGGTTCTTAAAGAT 363
QY 379 GAAAAGGCTAGATCACTCTTAGAGTCGAGCAGGAGGAGCGGGAATCTCTGGAGAGAGT 438
DB 364 GGAAGAGCACAACCTCTCTCTAGTGGGCTGGATGAAAGATCTCTGCAAT---CAGAGACT 420
QY 439 GGGCGGGGATATGAGCATGACATAGTAGTCTCTGTAATTTGTTGAGGAGAGAGAA 498
DB 421 GTTGAAGAGTGGGATGAGCAGCAGCAGCTCTGAAATTTGTTTCGAGGAGAGAGAAAG 480
QY 499 ACTATTTCAGCATCTAGAACCAAGAAATATGTTGCTGCAACGAGAAATATCTTCCGCTG 558
DB 481 AATTCAAGATATCAAGCTCAAGGAACAGTGTCTGCTTACAAGATAGGTGTTCTCTGTT 540
QY 559 CTCTCCCAAGGACCCAGTCTAGGTTTAAAGCCACAGCAGAAATGAGTCTGCAGCTACT 618
DB 541 CTCTCTAGAGGACCGGAGCAGGTTTACTCTCTCAAGAGCTTCAACAGGTTGTTGAAAAGCT 600
QY 619 CCAGTAGAGAGGAG---AAACAGCAGTAAATTTGTCACCAAGATCA 663
DB 601 GAACCAATGTTGGAGGCTACTGAAGCTGCAAGAAAGTTGCAGATCTGCTCCAGCTGTC 660
QY 664 GATCTCTCAAGAGTGGATTTCTACTGATCTGTTTAAACATGCTATCAATGATG 718
DB 661 TCTGCACCCAAAGTTGATTTTCTACTGACCTTNTCAACTTCTCTTCTCATGATG 715

RESULT 11
CO096428
LOCUS CO096428
DEFINITION GR_Ea19K21.r GR_Ea Gossypium raimondii cDNA clone GR_Ea19K21.3',
mRNA sequence.
ACCESSION CO096428
VERSION CO096428.1
KEYWORDS EST.
SOURCE Gossypium raimondii
825 bp mRNA linear EST 16-JUN-2004
GR_Ea19K21.r GR_Ea Gossypium raimondii cDNA clone GR_Ea19K21.3',
mRNA sequence.
CO096428
GI:48795114
```

[illegible]

```

AY596550      2173 bp      mRNA      linear      HTC 30-MAY-2004
LOCUS      Saccharum officinarum clone SCCREZ1001B01, complete sequence.
DEFINITION      AY596550
VERSION      AY596550.1 GI:47605274
KEYWORDS      HTC.
SOURCE      Saccharum officinarum
ORGANISM      Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE      1 (bases 1 to 2173)
AUTHORS      Oliveira,L.P., Carvalho,R.A., Drezza,A.L., Fregolente,M.C.,
Martini,I.J., Nascimento,B.O., Tsuneda,S.S., Rodrigues-Filho,P.C.,
Ulian,E.C., Nogueira,F.T.S., Vicentini,R., Felix,J.M. and
Menossi,M.
TITLE      Characterization of conserved hypothetical proteins from sugarcane
(Saccharum sp.)
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 2173)
AUTHORS      Oliveira,L.P., Carvalho,R.A., Drezza,A.L., Fregolente,M.C.,
Martini,I.J., Nascimento,B.O., Tsuneda,S.S., Rodrigues-Filho,P.C.,
Ulian,E.C., Nogueira,F.T.S., Vicentini,R., Felix,J.M. and
Menossi,M.
TITLE      Direct Submission
JOURNAL      Submitted (12-APR-2004) Departamento de Genetica e Evolucao,
Instituto de Biologia, Centro de Biologia Molecular e Engenharia
Genetica, Universidade Estadual de Campinas - Campinas (UNICAMP),
Campinas, SP 13083-970, Brazil
FEATURES      Location/Qualifiers
source      1..2173
              /organism="Saccharum officinarum"
              /mol_type="mRNA"
              /db_xref="taxon:4547"
              /clone="SCCREZ1001B01"

ORIGIN
Query Match      23.6%; Score 343.2; DB 3; Length 2173;
Best Local Similarity 61.2%; Pred. No. 7.7e-89;
Matches 656; Conservative 0; Mismatches 383; Indels 33; Gaps 5;

QY      1  ATGACGAGAAAGCAACGCTCTTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAA 60
DB      287  ATGACGAGAAAGGCTCGTTTCCAGAGCTCAACGCCAAGCAGACAGAGATTGGA 346
QY      61  GGGCTTCTTAACATCAGAGAAAGAGATGTCTGACTGCAAAACAAAAGGTCCAAGA 120
DB      347  AGTCTTCTCGGCTGCTGAGATAGAGATGCGCGACTGCAAGTCAAGGGTCTCTCGA 406
QY      121  TGGCTAGTGTAAATTAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT 180
DB      407  TGGCAAGTGTGAATCTAGGATCTTTATATGTATGATCATGTTCTGGAATTCATAGAAGC 466
QY      181  CTCGGGTGACATATCGAAGTTTCGATCTGCCACTCTGGACATCGCTCCCGAGCAG 240
DB      467  CTGGGGTGACATATCTAAGGTAAGATCTGCCACCTGGATACATCGCTCCAGACAA 526
QY      241  GTTCATTATPACAGTCAATGGGAAATGATAAAGCAAAATAGTTACTGGGAAGCAGCTA 300
DB      527  GTTGCAATTTTCAATCAATGGGAAACGAAAGCAAAATAGTATTGGGAAGCAGAGCTG 586
QY      301  CCCCCAACTATGATAGATTGGAAATTTTATACGTGCAAAAGTATGAGAGAAG 360
DB      587  CCTCCTAACTATGATAGATTGGGATGAGAAATTTTCATCGGTGCAAAATACGAGGA 646
QY      361  AGATGGGTTTCTAGAGGGAAGAGCTAGATCACTCTCTAGATCGAGCAGGACGGCG 420
DB      647  AGATGGGTACCGAGGAATGGAACATCAGACCTTCTCCGGTGTTCAGATGAGAAAGC 706
QY      421  AAACTGTGGAGAGAAGTGGCGGGATPATGAGCATGGACATAGTAGT--AGTCTGTAA 478
DB      707  CAAGAGTCTCGGCTAGTGCTTAATAGGAGTGGACATGTCATAGATCTTCAATTTGAGCA 766

```

```

479  ATTTGTTTGAGGAGAGAAACTATTTCAGCATCTAGAACAGAAATAATTTGTGCGAA 538
DB      767  AACCGTGGTTCCAGAGCTCTCCGAGCAAGTTGCACCTGTAGCTTCAAGGATACCTCT 826
QY      539  CGAGAAATAAA-TCTTCCCGTGCCTCCCAAGACCAGTCAGGTTATAAAGCCACAGCAG 597
DB      827  CAGGCATCACCTCAGCCCCCAAGGTTAGAACCAACCACTAGTTTCTTAAGGTGGTTTCA 886
QY      598  AAAATGGAGTCTGCAGTACTCAGTAGAGAGGGAGAAAACAGACGATAAATGTTGCACCA 657
DB      887  CAGCCACAGAAATCTCTCCAAAGTTGAGGCAACACCCCTTAAGTTGAGAGCCATCT 946
QY      658  GCATCAGATCTCTCCAAAGGTGGATTGCTACTGATCTGTTTAAACATGCTATCAATGGAT 717
DB      947  GTTGGCCCACTCTTAAAGTTGATTATGCTACTGATCTCTTTAAACATGTTTATCAATGGAT 1006
QY      718  GATTGACTACAAATACCTCAGAGGCACTCTCGGCGATACTCTCTCCGATGATACTCA 777
DB      1007  GGAACAAACAGAGAAAGAGTCAGAGTCATCTTCAAAC-----GATGATAATGCC 1054
QY      778  TGGGCTGGCTTTCAGTCTGCTGNAAGTGGTCAAAACGAGAGAGAAAATTTGCACAGCAAG 837
DB      1055  TGGGATGGCTTCCAGTCTGCAACCAACGATACCTAGCTCAGAGAAAAGATTCTGCCAAA 1114
QY      838  CTTGCTGAGAGCAGTTTCTCTCCAGCTTCATCTTTGAGCTTTGAGGATTGTTTAAAGGAC 897
DB      1115  CCAGCAGAAAGCAAGACCC-----AGTCAACATCTGGAATAGAGACTTATTTAAAGAC 1168
QY      898  ACACCTAATTTAACTCAACTCAACAGCACC-----AAAAGATGTGAAGGCGAT 945
DB      1169  TCACCGCTGTGTCAATATATCTCAGCTCCAGCTGTTTCCCAAGTAATGCGAAGAATGAT 1228
QY      946  ATCATGAGCTGTTTGAGAGAGAGATATAGTATCGCTTTTGCCATGCAATCAGCAACAG 1005
DB      1229  ATCATGAGTTGTTTGAAGAGTCCAATATGTTATCTCAATTTGCTGCCCATCAGCAGCAG 1288
QY      1006  GTTGCTATGCTGCTCAGCAGCAAGCCCTTTACATGCTGCGACGCAAAAGCTG 1057
DB      1289  CTGGCGTTATGTCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTTAAAGCTG 1340

CV197356      742 bp      mRNA      linear      EST 15-SEP-2004
CGF1003951_F06 Seed coat from mid-season walnut embryos collected
Aug 1 Juglans regia cDNA clone WSC0009_IF_F06 5', mRNA sequence.
CV197356
VERSION      CV197356.1 GI:52126193
KEYWORDS      EST.
SOURCE      Juglans regia (English walnut)
ORGANISM      Juglans regia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fagales; Juglandaceae; Juglans.
REFERENCE      1 (bases 1 to 742)
AUTHORS      Muir,R., Baek,J., Leslie,A., Cook,D. and Dandekar,A.
TITLE      Analysis of genes expressed in walnut seed coat tissue
JOURNAL      Unpublished (2004)
COMMENT      Contact: Abhaya Dandekar, PhD
              CAES Genome Facility
              UC Davis, Department of Pomology
              One Shields Ave, Davis, CA 95616, USA
              Tel: 530 752 7784
              Fax: 530 752 8502
              Email: amandandekar@ucdavis.edu
              Seq primer: WSCF-TCCGAGATCTGCACGAGC.
              Location/Qualifiers
FEATURES      1..742
source      /organism="Juglans regia"
              /mol_type="mRNA"
              /cultivar="Tulare"
              /db_xref="taxon:51240"

```


Wed Sep 28 08:10:50 2005

Qy	181	CTCGGGGTACACATATCGAAGGTTTCGATCTGCGCACTCTCGGACACATGGCTCCCGGACGAG	240
Db	479	CTGGGGGTGCACATATCTAAAGGTAAGATCTGCCACCCCTGGATACATGGCTGCCAGACAA	538
Qy	241	GTTCGATTTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTCGGGAAGCAGAGCTA	300
Db	539	GTTCGATTTTATTAATCAATGAGGAAACGAAAGCAAAATAGCTATTTGGGAAGCAGAGCTG	598
Qy	301	CCCCAAACTATGATAGAGTTGGAAATTTGAGAAATTTTATACGTGCAAAAGTATGAGAGAAG	360
Db	599	CCCTCTAACTACGATAGGGTTGGAATAGAGAAATTTTCATCCGTGCAAAATATAGGAGCAAG	658
Qy	361	AGATGGCTTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGAGTCGAGCAGGGAACGCGCG	420
Db	659	AGATGGGTACCAAGAAATGGAACATCAAAATCTTCGTGCTAGTGTTCCAGATAGAGAGAC	718
Qy	421	AAATCTGTGAGAGAAGTGGCGCGGGATATGAGCATGGACATAGTAGT--AGTCCTGTAA	478
Db	719	CAAGAGTCTCCGGCTAGTCTAATPAGAGTGGACATGGTCATAGATCTTCATTTGAGCAA	778
Qy	479	ATTTGTTTGGAGGAGAGAAACTATTTCCAGCATCTAGAAACAGAAATAATGTTGCTGCAA	538
Db	779	AACCGTGCCTTACCAGCTCTTCCGAGCAAAAGTTGCAATCGAGCTTCAAGGATATCCTCT	838
Qy	539	CGAGAAATAA--TCCTTCGGTGCTCCGCCAAGGACCCAGTCAGGTTATAAAGCCACACGAC	597
Db	839	CAGGCATCCTCAGCTCAGCCCCCGAGGTAGAAACACCACTTCCAAAGGTGGTTTACCTCCT	898
Qy	598	AAATGAGTCTGCAGTACTCTCAGTATAGAGGGAGAGAAACAGCAGTAAATGTTGCAACCA	657
Db	899	CAGTCACAGAAATCCCGCGCCAAAGTTGATGCAACACCGGCTAAAGTTTGAGAAGCCATCG	958
Qy	658	GCATCAGATCCTCCAAAGGTGGATTTTGTCTACTCATCTGTTTAACTAGCTATCAATGGAT	717
Db	959	GTTTACACACTCTCTAAAGTTGATATGCGCACTGATCTCTTTAACATGTTGTCAATGGAT	1018
Qy	718	GATTCGACTACAAATACCTCAGGCGAACTCTCTGGCGATACTCTGCGCATGATAACTCA	777
Db	1019	GGAAACAACAGAGAAAGAGTCAGCGTCA-----TCTTCAAAACGACGATAATGGC	1066
Qy	778	TGGCTGGCTTTTCACTCTGTGGAAGTGGTCAACGGCAGAGAAAATTTGCACAGCCAAG	837
Db	1067	TGGGATGGCTTCCAGTCTGCAACACCAAGTACCTAGTTTCAGAGAAAAAAGATCTGCCAAA	1126
Qy	838	CCTGCTCAGAGCAGCTTCTCTCCAGCTTTCATCTTCTGACTTTGAGGATTTGTTTAAAGGAC	897
Db	1127	CCAGCAGAAAGCAAGACCCAG-----TCTACATCTGGAATGGAAGACTTATTTAAGAC	1180
Qy	898	ACACTTAATTTTAAACAATCAACAGACCAAAAGATGTGAAGGC-----GAT	945
Db	1181	TCACAGCTGTGCCATTTATCCTCAGCTCCAGCTGTTTCCCAGTAAACGCTAAGAAATGAT	1240
Qy	946	ATCATGAGCTGTTTGGAGACGCAATATAGTATCGCTTTTGGCATGTCATCAGCAACAG	1005
Db	1241	ATCATAGTGTGTTTGGAGAGTCCAAATATGGTATCACTTTGCTGCGCCATCAACAGCAG	1300
Qy	1006	GTTGCTATGCTGCTCAGCAGCAAGCCCTTTTACATGGCTGAGCGGAAGCTG	1057
Db	1301	CTGCGGCTCATGTTCTCAGCAGCAGCTCTTCTTAATGGCTGCTGTTCAAGGTG	1352

Search completed: September 27, 2005, 06:18:03
Job time : 4465 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 02:39:09 ; Search time 779 Seconds
(without alignments)
11033.972 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaaacacacgt.....gaatgttcacaaacatgga 1452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a
-score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1452	100.0	1452	6	AAL48968
2	1452	100.0	1452	12	ADN72208
3	1452	100.0	1791	3	AAC48822
4	1450.4	99.9	1452	6	AAL48968
5	1450.4	99.9	1452	6	AAL48970
6	1447.2	99.7	1794	3	AAC38070
7	298.2	20.5	596	13	ACN48712
8	255.4	17.6	5580	6	AAL48972
9	237.2	16.3	371	3	AAC75657
10	237.2	16.3	371	6	ABN26360
11	227.4	15.7	313	6	ABL71943
12	199.6	13.7	270	10	ABX82114
13	182.4	12.6	408	3	AAC56661
14	127.2	8.8	290	6	ABL73349
15	107.2	7.4	587	6	ABQ98937
16	107.2	7.4	3117	4	AAC84451
17	105.6	7.3	799	4	AAH06694
18	105.6	7.3	2314	4	AAH17887
19	105.6	7.3	2314	8	ACD13426
20	98.4	6.8	1403	12	ADJ57942

21	94.8	6.5	2561	4	AAC84446	AAC84446 Nucleotid
22	94.2	6.5	667	4	AAI60996	AAI60996 Human pol
23	94.2	6.5	1447	6	ABS70464	ABS70464 Human bon
24	94.2	6.5	1976	4	AAI58627	AAI58627 Human pol
25	94.2	6.5	1976	5	ADQ98845	ADQ98845 DNA encod
26	94.2	6.5	1976	9	ADB48605	ADB48605 Novel hum
27	94.2	6.5	2457	8	ACD13427	ACD13427 Human DNA
28	94.2	6.5	2492	3	AACT76483	AACT76483 Human ORF
29	94.2	6.5	2618	13	ADP55646	ADP55646 Human PRO
30	94.2	6.5	2874	12	ADQ97937	ADQ97937 Human can
31	92.2	6.3	2541	2	AAQ04343	AAQ04343 Human sec
32	86.2	5.9	309	3	AAQ02721	AAQ02721 Human sec
33	82.6	5.7	1327	4	AAI59210	AAI59210 Human pol
34	82.6	5.7	1327	5	ADQ99434	ADQ99434 DNA encod
35	82.6	5.7	1327	9	ADB49194	ADB49194 Novel hum
36	82.6	5.7	2176	4	AAI60413	AAI60413 Human pol
37	82	5.6	1324	12	ADP22613	ADP22613 Sea-squir
38	82	5.6	1356	12	ADQ08635	ADQ08635 Ciona int
39	80.6	5.6	1274	3	AAC39358	AAC39358 Arabidops
40	80.4	5.5	507	10	ADG37966	ADG37966 Aspergill
41	80.4	5.5	1314	3	AAC66455	AAC66455 Human sec
42	80.4	5.5	2545	4	AAQ02552	AAQ02552 Human sec
43	79.4	5.5	555	13	ACN49796	ACN49796 Cotton pr
44	79.2	5.5	1775	12	ADP22611	ADP22611 Sea-squir
45	78.6	5.4	398	9	ACH29095	ACH29095 Human adu

ALIGNMENTS

RESULT 1
AAL48968
ID AAL48968 standard; DNA; 1452 BP.
XX
AC AAL48968;
XX
DT 28-OCT-2002 (first entry)
XX
DE A thaliana NEVERSHED protein coding sequence.

XX NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
KW GTP-binding protein domain; organ abscission; floral abscission; plant;
KW gene; ds.
XX Arabidopsis thaliana.
XX OS Arabidopsis thaliana.
XX FH Key Location/Qualifiers
XX CDS 1..1452
FT /*tag= a
FT /product= "NEVERSHED"
XX
PN WO200261042-A2.

XX PD 08-AUG-2002.
XX
XX 22-JAN-2002; 2002WO-US001938.
XX
XX 29-JAN-2001; 2001US-0264974P.
XX
XX (SALK) SALK INST BIOLOGICAL STUDIES.
XX (REGC) UNIV CALIFORNIA.
XX
XX Liljegren SJ, Becker JR, Yanofsky MF;
XX WPI; 2002-627475/67.
XX P-PSDB; AAC18759.
XX
XX New mutant plants comprising a modified ARF-GAP domain and having
XX decreased organ abscission, particularly floral abscission.
XX Claim 11; Page 44-46; 67pp; English.
XX The present invention relates to a mutant plant which has decreased organ
XX CC

CC abscission, comprising a nucleotide sequence having a modified ARF GAP
CC domain. The mutant plant is useful as a source of plants or progenies
CC having reduced or decreased floral abscission. The present sequence is
CC the Arabidopsis thaliana NEVERSHED protein coding sequence, which can be
XX mutated in plants of the invention

SQ Sequence 1452 BP; 461 A; 333 C; 329 G; 329 T; 0 U; 0 Other;

Query Match 100.0%; Score 1452; DB 6; Length 1452;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACGAGAGAGCAACGCTCTCTAAGGAGCTTAATGCCCGCCATAGAGAAGATCTCTGAA 60
DB 1 ATGAACGAGAGAGAGCAACGCTCTCTAAGGAGCTTAATGCCCGCCATAGAGAAGATCTCTGAA 60
QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGTGACTGTCAAAACAAAGAGTCCAAGA 120
DB 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGTGACTGTCAAAACAAAGAGTCCAAGA 120
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
DB 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
QY 181 CTCGGGTTACACATATCGAAGGTTGATCTGCACTCTGGACACATGGCTCCCGGAGCAG 240
DB 181 CTCGGGTTACACATATCGAAGGTTGATCTGCACTCTGGACACATGGCTCCCGGAGCAG 240
QY 241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB 241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
QY 301 CCCCAACTATGATAGAGTGGAAATGAGATTTTATACGTGCAAGTATGAAAGAG 360
DB 301 CCCCAACTATGATAGAGTGGAAATGAGATTTTATACGTGCAAGTATGAAAGAG 360
QY 361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAACGGCGG 420
DB 361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAACGGCGG 420
QY 421 AAATCTGTGGAGAGAGTGGCGGGGATATAGCATATGAGCATATAGTAGTAGTCTCTAAAT 480
DB 421 AAATCTGTGGAGAGAGTGGCGGGGATATAGCATATGAGCATATAGTAGTAGTCTCTAAAT 480
QY 481 TTGTTTGGAGAGAGGAAATCTATCCAGCATCTAGAACAGGAATAATGTTCTGCAACG 540
DB 481 TTGTTTGGAGAGAGGAAATCTATCCAGCATCTAGAACAGGAATAATGTTCTGCAACG 540
QY 541 AGAATAAATCTTCCCGTGTCTCCCAAGGACCCAGTCAAGTTATAAAGCCACAGCAGAAA 600
DB 541 AGAATAAATCTTCCCGTGTCTCCCAAGGACCCAGTCAAGTTATAAAGCCACAGCAGAAA 600
QY 601 ATGGAGTCTGACAGTACTCCAGTAGAGAGGGAGAAACAAAGCAGTAATGTTGCACACGA 660
DB 601 ATGGAGTCTGACAGTACTCCAGTAGAGAGGGAGAAACAAAGCAGTAATGTTGCACACGA 660
QY 661 TCAGATCTCCAAAGGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
DB 661 TCAGATCTCCAAAGGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
QY 721 TCGACTACAAATACCTCAGAGGCAACTCTCTGGCGATCTCTCTGCCGATGATACTCATGG 780
DB 721 TCGACTACAAATACCTCAGAGGCAACTCTCTGGCGATCTCTCTGCCGATGATACTCATGG 780
QY 781 GCTGGCTTTTCACTGTCTGGAAGTGTCAAAACGGCAGAGAAAATGTTCAAGCCAGCTT 840
DB 781 GCTGGCTTTTCACTGTCTGGAAGTGTCAAAACGGCAGAGAAAATGTTCAAGCCAGCTT 840
QY 841 GCTGAGAGAGAGTCTCTCCAGCTTCACTTTCTGACTTTGAGGATTTGTTTAAAGACACA 900
DB 841 GCTGAGAGAGAGTCTCTCCAGCTTCACTTTCTGACTTTGAGGATTTGTTTAAAGACACA 900
QY 901 CCTAATTTAACTCAACAGCACCACAAAGATGTGAAAGGGGATATCATGAGCCTGTTT 960

DB 901 CCTAATTTAACTCAACAGCACCACAAAGATGTGAAAGGGGATATCATGAGCCTGTTT 960
QY 961 GAGAAGAGCAATATAGTATGCGCTTTTGCCTATGATCAGCAACAGGTTGCTATGCTCGCT 1020
DB 961 GAGAAGAGCAATATAGTATGCGCTTTTGCCTATGATCAGCAACAGGTTGCTATGCTCGCT 1020
QY 1021 CAGCAGCAAGCCCTTTACATGGCTGCGAGCGAAAGCTGCTGAGGACACTCCAAACGGCGTG 1080
DB 1021 CAGCAGCAAGCCCTTTACATGGCTGCGAGCGAAAGCTGCTGAGGACACTCCAAACGGCGTG 1080
QY 1081 AATCAACAAGCTATTGTCTTAACTGCTTTTAACTGCTTAACTGCTTAACTGCTTAACTGCTT 1140
DB 1081 AATCAACAAGCTATTGTCTTAACTGCTTTTAACTGCTTAACTGCTTAACTGCTTAACTGCTT 1140
QY 1141 GGCTACAGATCCCGGAAATGACTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 GGCTACAGATCCCGGAAATGACTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 ATGCAAAACATGAATATGAACGCAAAACATGAACGCAACGCGGACCAACCGCAAGAGAAC 1260
DB 1201 ATGCAAAACATGAATATGAACGCAAAACATGAACGCAACGCGGACCAACCGCAAGAGAAC 1260
QY 1261 ACTCTAACAATCCCATCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
DB 1261 ACTCTAACAATCCCATCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
QY 1321 ATGACCCCAACTCAACCGGTAAACCTCAGTCAATCATCCGCAACCCCAACCAAGCACC 1380
DB 1321 ATGACCCCAACTCAACCGGTAAACCTCAGTCAATCATCCGCAACCCCAACCAAGCACC 1380
QY 1381 ACACCATCTTCACAATCAGCAAGAACTTTGATTTCTTCTCTTGATGGATGGAATGTTT 1440
DB 1381 ACACCATCTTCACAATCAGCAAGAACTTTGATTTCTTCTCTTGATGGATGGAATGTTT 1440
QY 1441 ACAAAACATTGA 1452
DB 1441 ACAAAACATTGA 1452
RESULT 2
ADN72208
ID ADN72208 standard; cDNA; 1452 BP.
XX
AC ADN72208;
XX
DT 15-JUL-2004 (first entry)
XX
DE Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 103.
XX
KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX
PN WO2004035798-A2.
XX
PD 29-APR-2004.
XX
PF 20-OCT-2003; 2003WO-EP011658.
XX
PR 18-OCT-2002; 2002EP-00079408.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Inze D, De Veyder L, Vlieghe K;
XX
DR WPI; 2004-348466/32.
XX
PT P-PSDB; ADN72209.
PT Altering plant characteristics, useful for producing plants for enzyme or pharmaceutical production comprises modifying in a plant, expression of

one or more nucleic acids and/or modifying level or activity of one or more proteins.

Claim 1: SEQ ID NO 103; 134pp; English.

This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up- or down-regulated in transgenic plants overexpressing the heterodimeric E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, plant architecture or physiology, altered endoreduplication, biochemistry, signal transduction, storage lipid mobilisation and/or altered photosynthesis, each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/or carbon metabolism or they function as transcription factors. This polynucleotide sequence is thale cress CDNA upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa transcription factor, given in an exemplification of the invention.

Sequence 1452 BP; 461 A; 333 C; 329 G; 329 T; 0 U; 0 Other;

Query Match 100.0%; Score 1452; DB 12; Length 1452;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
DB 1 ATGACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAAACAAAGGTCCAAGA 120
DB 61 GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAAACAAAGGTCCAAGA 120
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGTAATTTCTGGGATTCACAGGAGT 180
DB 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGTAATTTCTGGGATTCACAGGAGT 180
QY 181 CTCGGGTTACATATCGAAGGTTGATCTGCACTCTGGACACATGGCTCCCGCAGCAG 240
DB 181 CTCGGGTTACATATCGAAGGTTGATCTGCACTCTGGACACATGGCTCCCGCAGCAG 240
QY 241 GTTGCAATTTATACAGTCAATGGAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB 241 GTTGCAATTTATACAGTCAATGGAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
QY 301 CCCCCAACTATGATAGAGTTGGAATTTATACGTTTATACGTTGCAAAAGTATGAAGAGAG 360
DB 301 CCCCCAACTATGATAGAGTTGGAATTTATACGTTTATACGTTGCAAAAGTATGAAGAGAG 360
QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGTACCTCTAGAGTCAGCAGGAGGACGGCGG 420
DB 361 AGATGGGTTCTAGAGGGGAAAGGCTAGTACCTCTAGAGTCAGCAGGAGGACGGCGG 420
QY 421 AAATCTGTGAGAGAGTGGCGGGATATGACATGGACATAGTAGTAGTCTGTAAAT 480
DB 421 AAATCTGTGAGAGAGTGGCGGGATATGACATGGACATAGTAGTAGTCTGTAAAT 480
QY 481 TTGTTTGGAGAGGAGAAACTATTTCCAGCATCTAGAACCAAGAAATATGTTGCTGCAACG 540
DB 481 TTGTTTGGAGAGGAGAAACTATTTCCAGCATCTAGAACCAAGAAATATGTTGCTGCAACG 540
QY 541 AGAATTAATCTTCCCGTGGCTCCCGAGGACCCAGTCAGGTTATAAGCCACAGCAGAA 600
DB 541 AGAATTAATCTTCCCGTGGCTCCCGAGGACCCAGTCAGGTTATAAGCCACAGCAGAA 600
QY 601 ATGGAGCTCTGACGTACTCTCAGTAGAGGGGAGAAACAAGCAGTAAATGTTGCACCAGCA 660
DB 601 ATGGAGCTCTGACGTACTCTCAGTAGAGGGGAGAAACAAGCAGTAAATGTTGCACCAGCA 660

```

```

DB 601 ATGGAGCTCTGACGTACTCTCAGTAGAGGGGAGAAACAAGCAGTAAATGTTGCACCAGCA 660
QY 661 TCAGATCTCCAAAGGTGGATTTTGGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
DB 661 TCAGATCTCCAAAGGTGGATTTTGGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
QY 721 TCGACTACAAATACCTCAGAGGCAACTCTCTGGGATATCTCTGGGATATCAATGATGAT 780
DB 721 TCGACTACAAATACCTCAGAGGCAACTCTCTGGGATATCTCTGGGATATCAATGATGAT 780
QY 781 GCTGGCTTTCAGTCTGCTGGAGTGGTCAACGGCAGAGAAATTTGTCAAGCCAAAGCT 840
DB 781 GCTGGCTTTCAGTCTGCTGGAGTGGTCAACGGCAGAGAAATTTGTCAAGCCAAAGCT 840
QY 841 GCTCAGAGCAGTCTCTCTCCAGCTTTCATCTTCTGACTTTGAGGATTTGTTTAAAGACACA 900
DB 841 GCTCAGAGCAGTCTCTCTCCAGCTTTCATCTTCTGACTTTGAGGATTTGTTTAAAGACACA 900
QY 901 CCTAATTTAACTCAACAGGCACTCAACAGGCACTCAACAGGCACTCAACAGGCACTCAACAGGCACT 960
DB 901 CCTAATTTAACTCAACAGGCACTCAACAGGCACTCAACAGGCACTCAACAGGCACTCAACAGGCACT 960
QY 961 GAGAGCAGATATAGTATGCTCTTTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 GAGAGCAGATATAGTATGCTCTTTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 CAGCAGCAAGCCCTTTTATGCTGCTGAGCGAAAGCTGCTGAGGCACTCCAAACCGCGTG 1080
DB 1021 CAGCAGCAAGCCCTTTTATGCTGCTGAGCGAAAGCTGCTGAGGCACTCCAAACCGCGTG 1080
QY 1081 AATCAACAAGCTATTTGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCT 1140
DB 1081 AATCAACAAGCTATTTGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCT 1140
QY 1141 GGTACAGAGTCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GGTACAGAGTCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ATGCAAAACATGATATGAAACGCAAAATGAAACGCAAAATGAAACGCAAAATGAAACGCAAAATGAAAC 1260
DB 1201 ATGCAAAACATGATATGAAACGCAAAATGAAACGCAAAATGAAACGCAAAATGAAACGCAAAATGAAAC 1260
QY 1261 ACTCTACAAATCCCATCATCCAGTTTCTACAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320
DB 1261 ACTCTACAAATCCCATCATCCAGTTTCTACAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320
QY 1321 ATGACCCCAAACTCAACCGGTAACCTCAGTCATCATCGCAACCCCAACCCCAACCCCAACCCCAACCC 1380
DB 1321 ATGACCCCAAACTCAACCGGTAACCTCAGTCATCATCGCAACCCCAACCCCAACCCCAACCCCAACCC 1380
QY 1381 ACACCATCTTCACAATCAGGCAAAAGACTTTGATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
DB 1381 ACACCATCTTCACAATCAGGCAAAAGACTTTGATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
QY 1441 ACAAAACATTGA 1452
DB 1441 ACAAAACATTGA 1452

```

RESULT 3

AAC48822

ID AAC48822 standard; DNA; 1791 BP.

XX AAC48822;

AC AAC48822;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58899.

XX Hybridisation assay; genetic mapping; gene expression control.

KW protein identification; signal transduction pathway; metabolic pathway;

XX promoter; termination sequence; ss.

OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 18-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 04-MAY-1999; 99US-0132048P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 18-MAY-1999; 99US-0134370P.
XX 19-MAY-1999; 99US-0134768P.
XX 21-MAY-1999; 99US-0134941P.
XX 21-MAY-1999; 99US-0135124P.
XX 24-MAY-1999; 99US-0135353P.
XX 25-MAY-1999; 99US-0135629P.
XX 27-MAY-1999; 99US-0136021P.
XX 28-MAY-1999; 99US-0136392P.
XX 01-JUN-1999; 99US-0136782P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 08-JUN-1999; 99US-0137724P.
XX 10-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 14-JUN-1999; 99US-0138847P.
XX 16-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 17-JUN-1999; 99US-0139453P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 24-JUN-1999; 99US-0140354P.
XX 28-JUN-1999; 99US-0140695P.
XX 29-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140591P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145155P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149436P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

[illegible]

RESULT 4

AAAL48969
ID AAL48969 standard; DNA; 1452 BP.

XX AC
XX AAL48969;

DT 28-OCT-2002 (first entry)

DE A thaliana EMS mutated NEVERSHED protein coding sequence mutant nev-1.

XX NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
KW GTP-binding protein domain; organ abscission; floral abscission; plant;
KM mutant; nev-1; gene; ds.
XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

OS Synthetic.

FH Key Location/Qualifiers

FT CDS 1..1452

FT /*tag= a

FT /product= "nev-1"

XX WO200261042-A2.

PN 08-AUG-2002.

XX 22-JAN-2002; 2002WO-US001938.

XX 29-JAN-2001; 2001US-0264974P.

XX (SALK) SALK INST BIOLOGICAL STUDIES.

PA (REGC) UNIV CALIFORNIA.

XX Liljegren SU, Ecker JR, Yanofsky MF;

XX WPI; 2002-627475/67.

DR P-PSDB; AA018760.

XX New mutant plants comprising a modified ARF-GAP domain and having
PT decreased organ abscission, particularly floral abscission.

XX Claim 6; Page 48-50; 67pp; English.

XX The present invention relates to a mutant plant which has decreased organ
CC abscission, comprising a nucleotide sequence having a modified ARF GAP
CC domain. The mutant plant is useful as a source of plants or progenies
CC having reduced or decreased floral abscission. The present sequence is
CC the Arabidopsis thaliana NEVERSHED protein mutant coding sequence nev-1,
CC which can be used in plants of the invention

XX Sequence 1452 BP; 462 A; 333 C; 328 G; 329 T; 0 U; 0 Other;

Query Match 99.9%; Score 1450.4; DB 6; Length 1452;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACGAGAACCCACGCTCTTAAGAGCTTAATGCCCGGCATAGAAAGATTCTTGAA 60

DB 1 ATGAACGAGAACCCACGCTCTTAAGAGCTTAATGCCCGGCATAGAAAGATTCTTGAA 60

QY 61 GGGCTTTTAAACATCCAGAGACAGAGATGTGCTGACTGCACAAACAAAGGTCACAGA 120

DB 61 GGGCTTTTAAACATCCAGAGACAGAGATGTGCTGACTGCACAAACAAAGGTCACAGA 120

QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATCAATGTCTGGGATTACAGGAGT 180

DB 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATCAATGTCTGGGATTACAGGAGT 180

QY 181 CTCGGGGTACACATATCGAAGGTTTCGATCTGCCACTCTCGACACATGCTCCCGAGCAG 240

DB 181 CTCGGGGTACACATATCGAAGGTTTCGATCTGCCACTCTCGACACATGCTCCCGAGCAG 240

QY	241	GTTCATTATTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA	300
DB	241	GTTCATTATTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA	300
QY	301	CCCCAAAATATGATAGAGTTGGAAATTCAGAAATTTTATACGTGCAAAAGTATGAAGAGAAG	360
DB	301	CCCCAAAATATGATAGAGTTGGAAATTCAGAAATTTTATACGTGCAAAAGTATGAAGAGAAG	360
QY	361	AGATGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGAGAAACGGCGG	420
DB	361	AGATGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGAGAAACGGCGG	420
QY	421	AAATCTGTGGAGAGAAAGTGGGCGGGATATGAGCATGACATAGTAGTAGTCTCTTAAT	480
DB	421	AAATCTGTGGAGAGAAAGTGGGCGGGATATGAGCATGACATAGTAGTAGTCTCTTAAT	480
QY	481	TTGTTTGGAGAGAGAAAATTTTCCAGATCTAGAACAAAGAAATAATTTGTTCTGCAACG	540
DB	481	TTGTTTGGAGAGAGAAAATTTTCCAGATCTAGAACAAAGAAATAATTTGTTCTGCAACG	540
QY	541	AGATATAATCTTCCCGTCCCTCCCAAGGACCCAGTCTAGGTTTATAAGCCACAGCAGAA	600
DB	541	AGATATAATCTTCCCGTCCCTCCCAAGGACCCAGTCTAGGTTTATAAGCCACAGCAGAA	600
QY	601	ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAAACAAGCAGTAATTTGTCACACGCA	660
DB	601	ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAAACAAGCAGTAATTTGTCACACGCA	660
QY	661	TCAGATCTCCAAAAGTGGATTTTGTCTACTGATCTGTTTAAACATGCTATCAATGGATGAT	720
DB	661	TCAGATCTCCAAAAGTGGATTTTGTCTACTGATCTGTTTAAACATGCTATCAATGGATGAT	720
QY	721	TCGACTACAAATACCTCAGAGGCACTCTCGCGGATCTCTGCGGATGATTAACATCTG	780
DB	721	TCGACTACAAATACCTCAGAGGCACTCTCGCGGATCTCTGCGGATGATTAACATCTG	780
QY	781	GCTGCTTTTCACTGCTGGAAGTGGTCAAAAGGAGGAGAAAATTTGTCACGCCAAGCT	840
DB	781	GCTGCTTTTCACTGCTGGAAGTGGTCAAAAGGAGGAGAAAATTTGTCACGCCAAGCT	840
QY	841	GCTGAGCAGTCTCTCCAGCTTCAATCTGCTGAGGATTTGTTTAAAGCACA	900
DB	841	GCTGAGCAGTCTCTCCAGCTTCAATCTGCTGAGGATTTGTTTAAAGCACA	900
QY	901	CCTAATTTTAAACATCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	960
DB	901	CCTAATTTTAAACATCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	960
QY	961	GAGAGAGCAATATAGTATCGCTTTTTCCTATGCAATCAGCAACAGGTTGCTATGCTCGCT	1020
DB	961	GAGAGAGCAATATAGTATCGCTTTTTCCTATGCAATCAGCAACAGGTTGCTATGCTCGCT	1020
QY	1021	CAGCAGAACGCTTTTATGCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1080
DB	1021	CAGCAGAACGCTTTTATGCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1080
QY	1081	AATCAACAGCTATTGCTTAACGTAGCTTCTGCAAAATTTGGTCAAAACCCCGG	1140
DB	1081	AATCAACAGCTATTGCTTAACGTAGCTTCTGCAAAATTTGGTCAAAACCCCGG	1140
QY	1141	GGCTACACAGATCCCCGGGAATGACTAAACCCGCTAGGTGGTCAAGCTGATCTCCAGAACTT	1200
DB	1141	GGCTACACAGATCCCCGGGAATGACTAAACCCGCTAGGTGGTCAAGCTGATCTCCAGAACTT	1200
QY	1201	ATGCAAAACATGAATATGAACGCAAAATGAAACAGCAGACCCGCGCACACCGCAGAGAAC	1260
DB	1201	ATGCAAAACATGAATATGAACGCAAAATGAAACAGCAGACCCGCGCACACCGCAGAGAAC	1260
QY	1261	ACTCTACAAATCCCATCATCTCAGTTTCTTACAAATGGGTCAAGCTTAATCAAGTGAACGCT	1320
DB	1261	ACTCTACAAATCCCATCATCTCAGTTTCTTACAAATGGGTCAAGCTTAATCAAGTGAACGCT	1320

QY 1321 ATGACCCCAAACTCAACCGGTAAACCTCAGTCATCATCGCAACCAACCAACGACCC 1380
 DB |||||
 QY 1381 ACACCATCTTCAAAATCAGGCAAGACTTTGATTTCTCTCTTGATGATGGAATGTC 1440
 DB |||||
 QY 1441 ACAAAACATTGA 1452
 DB |||||
 QY 1441 ACAAAACATTGA 1452
 DB |||||
 RESULT 5
 ID AAL48970
 AC AAL48970 standard; DNA; 1452 BP.
 XX AAL48970;
 DT 28-OCT-2002 (first entry)
 DE A thaliana EMS mutated NEVERSHED protein coding sequence mutant nev-2.
 KW NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
 KW GTP-binding protein domain; organ abscission; floral abscission; plant;
 KW mutant; nev-2; gene; ds.
 XX Arabidopsis thaliana.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 1..594
 FT /*tag= a
 FT /product= "nev-2"
 XX WO200261042-A2.
 PN 08-AUG-2002.
 PD 22-JAN-2002; 2002WO-US001938.
 XX 29-JAN-2001; 2001US-0264974P.
 PR (SALK) SALK INST BIOLOGICAL STUDIES.
 PA (REGC) UNIV CALIFORNIA.
 PI Liljgren SJ, Ecker JR, Yanofsky MF;
 DR WPI: 2002-627475/67.
 DR P-PSDB; AAL08761.
 XX New mutant plants comprising a modified ARF-GAP domain and having
 PT decreased organ abscission, particularly floral abscission.
 XX Claim 6; Page 52-53; 67pp; English.
 CC The present invention relates to a mutant plant which has decreased organ
 CC abscission, comprising a nucleotide sequence having a modified ARF GAP
 CC domain. The mutant plant is useful as a source of plants or progenies
 CC having reduced or decreased floral abscission. The present sequence is
 CC the Arabidopsis thaliana NEVERSHED protein mutant coding sequence nev-2,
 CC which can be used in plants of the invention
 XX Sequence 1452 BP; 461 A; 332 C; 329 G; 330 T; 0 U; 0 Other;
 SQ
 Query Match 99.9%; Score 1450.4; DB 6; Length 1452;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGACGAGAGAGCCAAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
 DB |||||
 1 ATGACGAGAGAGCCAAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60

QY 61 GGGCTTTCTTAATCAATCCAGAGAACAGAGAAATGTGCTGATCTGCAAAACAAAAGGTCCACGA 120
 DB |||||
 QY 61 GGGCTTTCTTAATCAATCCAGAGAACAGAGAAATGTGCTGATCTGCAAAACAAAAGGTCCACGA 120
 DB |||||
 QY 121 TGGGCTAGTGTAAATTAAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAT 180
 DB |||||
 QY 121 TGGGCTAGTGTAAATTAAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAT 180
 DB |||||
 QY 181 CTGGGGGTACACATATCGAAGGTTGATCTGCACTCTGGACACATGGCTCCCGAGCAG 240
 DB |||||
 QY 181 CTGGGGGTACACATATCGAAGGTTGATCTGCACTCTGGACACATGGCTCCCGAGCAG 240
 DB |||||
 QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
 DB |||||
 QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
 DB |||||
 QY 301 CCCCCAAACTATAGATAGAGTTGGAAATTCAGAAATTTTATACGTGCAAGTATCAAGAGAG 360
 DB |||||
 QY 301 CCCCCAAACTATAGATAGAGTTGGAAATTTGAAATTTTATACGTGCAAGTATCAAGAGAG 360
 DB |||||
 QY 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAAACGGCG 420
 DB |||||
 QY 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAAACGGCG 420
 DB |||||
 QY 421 AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATAGGACATAGTAGTAGTCTCTGTAAT 480
 DB |||||
 QY 421 AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATAGGACATAGTAGTAGTCTCTGTAAT 480
 DB |||||
 QY 481 TTGTTTGGAGAGAGAAAATCTATTTCCAGCATCTAGAACAAAGAAATTAATGTTTCTGCAACG 540
 DB |||||
 QY 481 TTGTTTGGAGAGAGAAAATCTATTTCCAGCATCTAGAACAAAGAAATTAATGTTTCTGCAACG 540
 DB |||||
 QY 541 AGATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCAGGTTTATAAAGCCACAGCAGAAA 600
 DB |||||
 QY 541 AGATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCAGGTTTATAAAGCCATAGCAGAAA 600
 DB |||||
 QY 601 ATGGAGTCTGCAGCTACTCTCAGTAGAGAGGAGAAACAAAGCAGTAAATGTTGCAACGCA 660
 DB |||||
 QY 601 ATGGAGTCTGCAGCTACTCTCAGTAGAGAGGAGAAACAAAGCAGTAAATGTTGCAACGCA 660
 DB |||||
 QY 661 TCAGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
 DB |||||
 QY 661 TCAGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
 DB |||||
 QY 721 TCGACTACAAATACCTCAGAGGCAACTCTCGCGATACTCTCGCGATGATTAATCTCATG 780
 DB |||||
 QY 721 TCGACTACAAATACCTCAGAGGCAACTCTCGCGATACTCTCGCGATGATTAATCTCATG 780
 DB |||||
 QY 781 GCTGGCTTTTCAGTCTGTGGAAGTGTCAAACGGCAGAGAAAATTTGTCAAGCCAGCCT 840
 DB |||||
 QY 781 GCTGGCTTTTCAGTCTGTGGAAGTGTCAAACGGCAGAGAAAATTTGTCAAGCCAGCCT 840
 DB |||||
 QY 841 GCTGAGAGAGTCTCCTCCAGCTTCACTTCTGATCTTTGAGGATTTGTTTAAAGCAGCA 900
 DB |||||
 QY 841 GCTGAGAGAGTCTCCTCCAGCTTCACTTCTGATCTTTGAGGATTTGTTTAAAGCAGCA 900
 DB |||||
 QY 901 CCTAATTTAAACACTCAACAAAGCATGTAAGAGGCGATATCATAGAGCCTGTTT 960
 DB |||||
 QY 901 CCTAATTTAAACACTCAACAAAGCATGTAAGAGGCGATATCATAGAGCCTGTTT 960
 DB |||||
 QY 961 GAGAAAGAGAAATATAGTATCGCTTTTCCATGATCAGCAACAGGTTGCTATGCTCGCT 1020
 DB |||||
 QY 961 GAGAAAGAGAAATATAGTATCGCTTTTCCATGATCAGCAACAGGTTGCTATGCTCGCT 1020
 DB |||||
 QY 1021 CAGCAGAGAGCCTTTTACATGCTGCGGAAAGTGTGGAGGCACTCCAAACGGCGTG 1080
 DB |||||
 QY 1021 CAGCAGAGAGCCTTTTACATGCTGCGGAAAGTGTGGAGGCACTCCAAACGGCGTG 1080
 DB |||||
 QY 1081 AATCAACAGCTATTGCTTAATGCTCTTAAACGTAAGTCTTCTGCAAAATTTGGTCAAAACCCCGC 1140
 DB |||||
 QY 1081 AATCAACAGCTATTGCTTAATGCTCTTAAACGTAAGTCTTCTGCAAAATTTGGTCAAAACCCCGC 1140
 DB |||||
 QY 1141 GGCTACCAGATCCCGGAAATGATTAACCCCGTAGGTGCTCAAGCTGATCTCCAGAAACTT 1200

Db 1141 |||||GGCTACCCAGATCCCGGAATGACTTAACCCCGTAGGTGGTCAAGCTGATCTCCAGAAACTT 1200
Qy 1201 ATGCAAAACATGAATATGACCGCAACACGACGAGACCGGACCAACCGCAAGGAC 1260
Db 1201 ATGCAAAACATGAATATGACCGCAACACGACGAGACCGGACCAACCGCAAGGAC 1260
Qy 1261 ACTCTACAAATACCCATCATCCAGTTTCTTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
Db 1261 ACTCTACAAATACCCATCATCCAGTTTCTTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
Qy 1321 ATGACCCCAACTCAACCGGTAAACCTCAGTCATCATCCGCAACCCCAACCAACGACACC 1380
Db 1321 ATGACCCCAACTCAACCGGTAAACCTCAGTCATCATCCGCAACCCCAACCAACGACACC 1380
Qy 1381 ACACCATCTTCAATCAGCGCAAGACTTTGATTTCTCTTCTGATGATGGAATGTC 1440
Db 1381 ACACCATCTTCAATCAGCGCAAGACTTTGATTTCTCTTCTGATGATGGAATGTC 1440
Qy 1441 ACAAAACATTGA 1452
Db 1441 ACAAAACATTGA 1452

RESULT 6
AAC38070
ID AAC38070 standard; DNA; 1794 BP.
XX AC AAC38070;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19684.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic pathway;
XX KW promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 16-APR-1999; 99US-0128714P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 28-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132407P.
XX PR 04-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 11-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.	Query Match		
PR	05-AUG-1999;	99US-0147260P.	Best Local Similarity 99.7%; Score 1447.2; DB 3; Length 1794;		
PR	06-AUG-1999;	99US-0147303P.	Matches 1449; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
PR	06-AUG-1999;	99US-0147416P.			
PR	06-AUG-1999;	99US-0147493P.	QY	1	ATGAACGAGAAAGCCACGCTCTTAAGGAGCTTAATGCCGCCCATAGAAAGATCTTTCGA 60
PR	09-AUG-1999;	99US-0147935P.	Db	167	ATGAACGAGAAAGCCAAATGTCTTAAGGAGCTTAATGCCGCCCATAGAAAGATCTTTCGA 226
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.	QY	61	GGGCTTCTTAACATCCAGAGAACAGAGAAATGCTGACTGCAAAACAAAAGGTCCAAGA 120
PR	12-AUG-1999;	99US-0148341P.	Db	227	GGGCTTCTTAACATCCAGAGAACAGAGAAATGCTGACTGCAAAACAAAAGGTCCAAGA 286
PR	13-AUG-1999;	99US-0148565P.			
PR	16-AUG-1999;	99US-0149368P.	QY	121	TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
PR	17-AUG-1999;	99US-0149175P.	Db	287	TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 346
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.	QY	181	CTCGGGGTACACATATCGAAGGTTCCGACTCTGCCACTCTGGACACATGCTCCCCGAGCAG 240
PR	20-AUG-1999;	99US-0149723P.	Db	347	CTCGGGGTACACATATCGAAGGTTCCGACTCTGCCACTCTGGACACATGCTCCCCGAGCAG 406
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.	QY	241	GTTGCATTTTACAGTCAATCGGAAATGATAAGCAATAGTTACTTGGGAAGCAGAGCTTA 300
PR	25-AUG-1999;	99US-0150566P.	Db	407	GTTGCATTTTACAGTCAATCGGAAATGATAAGCAATAGTTACTTGGGAAGCAGAGCTTA 466
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.	QY	301	CCCCCAACTATGATAGATTGGAAATGAGAAATTTTATACGTGCAAAAGTATGAAGAGAG 360
PR	27-AUG-1999;	99US-0151080P.	Db	467	CCCCCAACTATGATAGATTGGAAATGAGAAATTTTATACGTGCAAAAGTATGAAGAGAG 526
PR	30-AUG-1999;	99US-0151303P.			
PR	01-SEP-1999;	99US-0151438P.	QY	361	AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTCGAGCAGAGAAAGCGGG 420
PR	07-SEP-1999;	99US-0152363P.	Db	527	AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTCGAGCAGAGAAAGCGGG 586
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.	QY	421	AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATGAGCATAGTAGTAGTCTCTTAAT 480
PR	16-SEP-1999;	99US-0154018P.	Db	587	AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATGAGCATAGTAGTAGTCTCTTAAT 646
PR	20-SEP-1999;	99US-0154039P.			
PR	22-SEP-1999;	99US-0155139P.	QY	481	TGTTTGTAGGAGAGGAAAATTAATTCAGCATCTAGAACAGAAATAATTTGTTCTGCAACG 540
PR	23-SEP-1999;	99US-0155486P.	Db	647	TGTTTGTAGGAGAGGAAAATTAATTCAGCATCTAGAACAGAAATAATTTGTTCTGCAACG 706
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.	QY	541	AGAATAAATCTTCCGTCGCTCCCAAGGACCCAGTCCAGTCTAGGTTTAAAGCCAGCAGAAA 600
PR	29-SEP-1999;	99US-0156596P.	Db	707	AGAATAAATCTTCCGTCGCTCCCAAGGACCCAGTCCAGTCTAGGTTTAAAGCCAGCAGAAA 766
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.	QY	601	ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGGAGAAACAGCAGTAAATTTGTTCCACGAGCA 660
PR	06-OCT-1999;	99US-0157865P.	Db	767	ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGGAGAAACAGCAGTAAATTTGTTCCACGAGCA 826
PR	07-OCT-1999;	99US-0158029P.			
PR	12-OCT-1999;	99US-0158232P.	QY	661	TCAGATCTCTCAAAAGGTGGATTTTGTACTGATCTGTTTAAACATGCTCTCAATGGATGAT 720
PR	13-OCT-1999;	99US-0159293P.	Db	827	TCAGATCTCTCAAAAGGTGGATTTTGTACTGATCTGTTTAAACATGCTCTCAATGGATGAT 886
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.	QY	721	TCGACTACAAATACTCAGAGGCAACTCTCTGGCGATCTCTCTGGCGATGATTAATCTCATGG 780
PR	14-OCT-1999;	99US-0159330P.	Db	887	TCGACTACAAATACTCAGAGGCAACTCTCTGGCGATCTCTCTGGCGATGATTAATCTCATGG 946
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.	QY	781	GCTGGCTTTCAAGTCTGCTGGAGTGGTCAAAAGGAGAGAAATTTGTACAGCCCAAGCCT 840
PR	18-OCT-1999;	99US-0159638P.	Db	947	GCTGGCTTTCAAGTCTGCTGGAGTGGTCAAAAGGAGAGAAATTTGTACAGCCCAAGCCT 1006
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.	QY	841	GCTGAGAGCAGTTCTCTCCAGCTTCACTCTTCTGACTTTGAGGATTTGTTTAAAGGACACA 900
PR	21-OCT-1999;	99US-0160768P.	Db	1007	GCTGAGAGCAGTTCTCTCCAGCTTCACTCTTCTGACTTTGAGGATTTGTTTAAAGGACACA 1066
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.	QY	901	CCTAATTTTAACTCAACAGCACCAAAAGATGTGAAAGGCGATATCATGAGCCTGTTT 960
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160989P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			

RESULT 8
 AAL48972/c
 ID AAL48972 standard; DNA; 5580 BP.
 XX
 AC AAL48972;
 XX
 DT 28-OCT-2002 (first entry)
 XX
 DE A thaliana NEVERSHED protein gene.
 XX
 KW NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
 KW GTP-binding protein domain; organ abscission; floral abscission; plant;
 KW gene; db.
 OS Arabidopsis thaliana.
 XX
 PN WO200261042-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 22-JAN-2002; 2002WO-US001938.
 XX
 PR 29-JAN-2001; 2001US-0264974P.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Liljegren SJ, Ecker JR, Yanofsky MF;
 DR WPI; 2002-627475/67.
 XX
 PT New mutant plants comprising a modified ARF-GAP domain and having
 PT decreased organ abscission, particularly floral abscission.
 PS Example 2; Page 54-56; 67pp; English.
 CC The present invention relates to a mutant plant which has decreased organ
 CC abscission, comprising a nucleotide sequence having a modified ARF GAP
 CC domain. The mutant plant is useful as a source of plants or progenies
 CC having reduced or decreased floral abscission. The present sequence is
 CC the Arabidopsis thaliana NEVERSHED protein coding sequence, which can be
 CC mutated in plants of the invention
 SQ Sequence 5580 BP; 1821 A; 1010 C; 1034 G; 1715 T; 0 U; 0 Other;
 Query Match 17.6%; Score 255.4; DB 6; Length 5580;
 Best Local Similarity 60.1%; Pred. No. 1.8e-66;
 Matches 877; Conservative 0; Mismatches 1; Indels 581; Gaps 4;
 QY 575 GTCAGGTTATAAGCCACAGCAGGAAATGGAGTCTGCAGTACTCCAGTAGAGGGGAGA 634
 DB 1482 GACAGGTTATAAGCCACAGCAGGAAATGGAGTCTGCAGTACTCCAGTAGAGGGGAGA 1423
 QY 635 AACACAGCTAATGTTGCACAGCATCAGATCTCCAAAGTGGATTGTTCTACTGATC 694
 DB 1422 AACACAGCTAATGTTGCACAGCATCAGATCTCCAAAGTGGATTGTTCTACTGATC 1363
 QY 695 TGTTTAACTGCTATCAATGGATGATTCGACTACAAATACCTCAGAGGCAACTCTGGCG 754
 DB 1362 TGTTTAACTGCTATCAATGGATGATTCGACTACAAATACCTCAGAGGCAACTCTGGCG 1303
 QY 755 ATACTCTCCGATGATACTATGGCGTGGCTTTGAGT----- 793
 DB 1302 ATACTCTCCGATGATACTATGGCGTGGCTTTGAGTGTATCTATCTAGCATGT 1243
 QY 794 ----- 793
 DB 1242 GATAATCTAACATCGTGGCACATTTTTCCTATCCCTGTTAACTTTATGCAATGTTCTT 1183
 QY 794 -----CTGCTGG 800
 DB 1182 CTTTGAATAATGCTTTGGAGTAAATAGATTAATTTCTGAATCTTAAACTGTAGCTGCTGG 1123

QY 801 AAGTGGTCAAAACGCGAGAGAAAATTTGTCAAGCCAAAGCTGCTGAGAGCAGTTCTCTCC 860
 DB 1122 AAGTGGTCAAAACGCGAGAGAAAATTTGTCAAGCCAAAGCTGCTGAGAGCAGTTCTCTCC 1063
 QY 861 AGCTTTCATCTTCTGACTTTGAGGATTTGTTTAAAGGACACACTAATTTAACTCAACA 920
 DB 1062 AGCTTTCATCTTCTGACTTTGAGGATTTGTTTAAAGGACACACTAATTTAACTCAACA 1003
 QY 921 AGCACCACAAAGATGTGAAGGGGATATCATGAGCTGTTTGAGA----- 964
 DB 1002 AGCACCACAAAGATGTGAAGGGGATATCATGAGCTGTTTGAGAAGGTAATATTAGTT 943
 QY 965 ----- 964
 DB 942 AATTATATACGTTCTCAGATTTCACTGCTTTGTTGATATTCAAGCTTTTCCAAAGTTTCGTT 883
 QY 965 ----- 964
 DB 882 TTTGTATCATTTAATTAACTTACATTTTCAAAATATAGCGAATGATATGATTTGTGTTTGT 823
 QY 965 ----- 964
 DB 822 TATCATTACTATGATTTCTTCAACCGGACATTTGGAAACACGAGAGAGATTCAAGTG 763
 QY 965 -----AGACGAATATAGTATCGCCTTTTGGCCATGCA 995
 DB 762 ATATTAACCTCTGTGTGTTCTATTGTTTACAGACGATATAGTATCGCCTTTTGGCCATGCA 703
 QY 996 TCAGCAACAGGTTGCTATGCTGCTCAGCAGCAAGCCCTTTACATGGTGCAGCGAAAGC 1055
 DB 702 TCAGCAACAGGTTGCTATGCTGCTCAGCAGCAAGCCCTTTACATGGTGCAGCGAAAGC 643
 QY 1056 TGCTGGAGGCACTCCAAACGCGGTGAATCAACAGCTATTGCTAATGCTCTTAAAGTACG 1115
 DB 642 TGCTGGAGGCACTCCAAACGCGGTGAATCAACAGCTATTGCTAATGCTCTTAAAGTACG 583
 QY 1116 TTCTGCAAAATTTGGTCAAAACCGCGGCTACAGATCCCGGAATGACTTAACCCGCTAGG 1175
 DB 582 TTCTGCAAAATTTGGTCAAAACCGCGGCTACAGATCCCGGAATGACTTAACCCGCTAGG 523
 QY 1176 TGGTCAAGCTGATCTCCAGAACTTATGCA----- 1206
 DB 522 TGGTCAAGCTGATCTCCAGAACTTATGCAAGTAACTCTCCCTTTACTTTTTTAAAGC 463
 QY 1207 ----- 1206
 DB 462 AAGATCAATATTTACATTTGACCAAAAGTTCTTGCAAAATTTCTCACTAAATTTCTGCTCTCTG 403
 QY 1207 -----AACATGAATATGAACGCAAAACATGAACACGAGACCGCGCAC 1246
 DB 402 TTACATCGCCATTAATTTTAGAACATGATATGAACGCAAAACATGAACACGAGACCGCGCAC 343
 QY 1247 AACCGCAAGAGAACACTCTACAAATACCCATCATC----- 1280
 DB 342 AACCGCAAGAGAACACTCTACAAATACCCATCATCAGGATATGCACACAAACTGCATCTT 283
 QY 1281 ----- 1280
 DB 282 GTTTGTTTCTTCTGAAACATCATAAAGTTATCAAAAGTCTTTAACTTAACCTAATCTCTCA 223
 QY 1281 -----CAGTTTCTTACACAATGGGTCAAGCTTAATCAAGT 1313
 DB 222 TATGATCTTTGGTGTGTTTAAATGAGTTTCTACACAATGGGTCAAGCTTAATCAAGT 163
 QY 1314 GAACGGTATGACCCCAAACTCAACCGGTAACCTCAGTGCATCATCGGACACCAACCAAC 1373
 DB 162 GAACGGTATGACCCCAAACTCAACCGGTAACCTCAGTGCATCATCGGACACCAACCAAC 103
 QY 1374 AAGCACCAACCATCTTTCACATCAGGCAAGACTTTGATTTCTCTTCTCTGATGATGG 1433
 DB 102 AAGCACCAACCATCTTTCACATCAGGCAAGACTTTGATTTCTCTTCTCTGATGATGG 43

QY 1434 AATGTTACAAACATTTGA 1452
 DB |||||
 42 AATGTTACAAACATTTGA 24

RESULT 9
 AAC75657
 ID AAC75657 standard; cDNA; 371 BP.
 XX
 AC AAC75657;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF1212 polynucleotide sequence SEQ ID NO:2423.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US008621.
 XX
 PR 31-MAR-1999; 99US-0127607P.
 PR 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 XX WPI; 2000-602362/57.
 DR P-PSDB; AAB41448.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 5; Page 1753-1754; 5507pp; English.
 XX
 CC AAC74446 to AAC75606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritis; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancer, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune

disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 enhance coagulation; to inhibit thrombosis; and as a contraceptive

Sequence 371 BP; 113 A; 74 C; 102 G; 82 T; 0 U; 0 Other;

Query Match 16.3%; Score 237.2; DB 3; Length 371;
 Best Local Similarity 81.4%; Pred. No. 1.5e-61;
 Matches 275; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAAGCCACGCTCTTAAGGAGCTTAATGCCCCCATAGAAAGATTCTTGAA 60
 DB |||||
 34 ATGAACGAGAAAGCGCTCGTCTCAAGGAGCTCAACGCCAAGCACACAGAGATATTGGA 93
 QY 61 GGGCTTCTTAACATCCAGAGAACAGAGAAATGCTGCTGCAAAACAAAGGTCCAAGA 120
 DB |||||
 94 GGTCTTCTACGGCATCTCGAATAGAGAAATGCGCAGACTGCAAGTCAAGGGTCTCTGA 153
 QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCAATGTTCTGGGATTCACAGGAGT 180
 DB |||||
 154 TGGCAAGTGTGATCTAGGTATCTTTATGATGATGATGATCTCTGGCATTTCATAGAAGC 213
 QY 181 CTCGGGGTACACATATCAAGGTTTCGATCTGCCACTCTTCGACATATGCTCCCGAGAG 240
 DB |||||
 214 CTGGGGGTGCACATATCTAAGGTAAAGATCTGCCACCTCTGGATACATGCTGCGCAGAGCA 273
 QY 241 GTTCGATTTATACAGTCAATGGGAAATGATAAAGCAAAATAGTCTTCTGGGAAAGCAGAGCTA 300
 DB |||||
 274 GTTCGATTTATCAATCAATGGGAAACGAAAGCAAAATAGTCTTCTGGGAAAGCAGAGCTG 333
 QY 301 CCCCCAACTATGATAGAGTTGGAAATTCGAAATTTAT 338
 DB |||||
 334 CCTCCTAACTACGATAGGTTGGAAATAGAGAAATTTGAT 371

RESULT 10
 AEN26360
 ID AEN26360 standard; cDNA; 371 BP.
 XX
 AC AEN26360;
 XX
 DT 24-JUN-2002 (first entry)
 XX
 DE Human ORFX polynucleotide sequence SEQ ID NO:21197.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.
 OS
 XX Homo sapiens.
 XX
 PN WO200192523-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US010836.
 XX
 PR 30-MAY-2000; 2000US-0206132P.
 PR 29-AUG-2000; 2000US-0228716P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach MD;
 XX
 XX WPI; 2002-106308/14.
 DR P-PSDB; ABE10608.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 21197; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification)). ABLN15762 to ABLN27252 encode the human ORFX
CC proteins given in ABLN0010 to ABLN1500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 371 BP; 113 A; 74 C; 102 G; 82 T; 0 U; 0 Other;
XX
Query Match 16.3%; Score 237.2; DB 6; Length 371;
Best Local Similarity 81.4%; Pred. No. 1.5e-61;
Matches 275; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
XX
QY 1 ATGACGAGAACCAACGCTCTTAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAA 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 34 ATGAACGAGAAGCGTCCGTCTCAAGGAGCTCAAGCCACGACGACGAGAGATATTGNA 93
QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAGA 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 94 GGTCTTCTACGGCATCTCGAATAGAGAAATGCCAGACTGCAAGTCAAAAGGTCTCTCGA 153
QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATCATGCTGCTGGGATTCACAGGACT 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 154 TGGGCAAGTGTGAATCTAGGTATCTTTATATGCAATGACATGTCTGGCAATTCATAGAAGC 213
QY 181 CTCGGGGTACACATATCGAAGGTTCGATCTGCCACTCTGCACATGCGCTCCCGAGCAG 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 214 CTGGGGGTGGACATATCTTAGGTAAAGATCTGCCACCTTGATACATGGTTCGACAGCA 273
QY 241 GTTGCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTTACTTGGGAAGCAGCTA 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 274 GTTGCAATTTATCAATCAATGGGAACGAAAGCAAAATAGCTATTGGGAAGCAGAGCTG 333
QY 301 CCCCCAACTATGATGAGTTGGAATTTAT 338
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 334 CCTCCTAACTACGATAGGTGGAAATAGAGAATTTGAT 371
XX
RESULT 11
ABL71943
ID 'ABL71943 standard; cDNA; 313 BP.
XX
AC ABL71943;
XX
XX 14-MAY-2002 (first entry)
XX
XX Corn tassell-derived polynucleotide (cdps) SEQ ID NO:1317.
XX
XX Corn; corn tassell-derived polynucleotide; cdps; hybrid breeding; CDPs;
XX inheritance; characteristic; growth; development; disease resistance;
KW

environmental adaptability; quality; yield; molecular marker;
multigene trait; plant breeding; corn tassell; gene; ss.
Zea mays.
XX
XX US2001051335-A1.
XX
XX 13-DEC-2001.
XX
XX 16-APR-1999; 99US-00294093.
XX
XX 21-APR-1998; 98US-0082567P.
XX
XX (LALG// LALGUDI R V.
XX (ITOL// ITO L Y.
XX (SHER// SHERMAN B K.
XX
XX Laigudi RV, Ito LY, Sherman BK;
XX WPI; 2002-163647/21.
XX
XX Novel purified corn tassell-derived polynucleotide useful for determining
XX altered gene expression, to recover regulatory elements and to follow
XX inheritance of desirable characteristics through hybrid breeding
XX programs.
XX
XX Claim 1; SEQ ID NO 1317; 201pp; English.
XX
XX The present sequence describes a purified corn tassell-derived
XX polynucleotide sequence (cdps) comprising a nucleic acid sequence
XX selected from those given in ABL70627 to ABL76833. The cdps sequences
XX encode corn tassell-derived polypeptides (CDPs). The cdps sequences (I)
XX can be used for determining altered gene expression, to recover
XX regulatory elements and to follow inheritance of desirable
XX characteristics through hybrid breeding programs. (I) are also useful in
XX the evaluation, and alteration of desired characteristics associated with
XX growth and development, disease resistance, environmental adaptability,
XX quality and yield, and as molecular markers for studying inheritance of
XX multigene traits in a plant breeding program. (I) can be used to produce
XX a tassell-specific profile of gene transcription, a transcript image, to
XX clone regulatory elements for use in transformation vectors, to express a
XX polypeptide, to identify, isolate or extend identical or related corn
XX tassell nucleic acid sequences from DNA libraries, in nucleic acid
XX hybridisation or amplification technologies, as query sequences to
XX determine homology of known sequences, as probe for use in Southern or
XX Northern hybridisation, and to identify the presence of and/or to
XX determine the degree of similarity between two (or more) nucleic acid
XX sequences
XX
XX Sequence 313 BP; 98 A; 59 C; 82 G; 74 T; 0 U; 0 Other;
XX
Query Match 15.7%; Score 227.4; DB 6; Length 313;
Best Local Similarity 83.5%; Pred. No. 1.4e-58;
Matches 258; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
XX
QY 60 AGGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCACAAACAAAAGGTCCAAG 119
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1 AGGTCTTCTACGGCATCTCGAATAGAGATGCGCAGACTGCAAGTCAAAAGGTCTCTCG 60
QY 120 ATGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCAATGTTCTGGGATTCACAGGAG 179
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 61 ATGGGCAAGTGTGAATCTAGGTATCTTTATATGCAATGACATGTTCTGGCAATTCATAGAAG 120
QY 180 TCTCGGGGTACACATATCGAAGTTTCGATCTGCCACTCTGGACACATGCTCCCCGAGCA 239
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 121 CCTGGGGTGGACATATCTTAAGGTAAAGATCTGCCACCTTGATACATGCTGCCAGCA 180
QY 240 GGTTCATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTTACTTGGGAGCAGAGCT 299
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 181 AGTTCATTTATTCAATCAATGGAACGAAAGCAAAATAGCTATTGGGAAGCAGAGCT 240
QY 300 ACCCCCAACTATGATGAGTTGGAATTTATACGTCGCAAACTATGCAAGAA 359
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

241 GCCTCTAAGTACGATAGGTTGGATAGAGAAATTTTCCTCGTCAAAATATGAGGACAA 300

360 GAGATGGGT 368

301 GAGATGGGT 309

RESULT 12

ABX82114

ID ABX82114 standard; cDNA; 270 BP.

XX ABX82114;

24-APR-2003 (first entry)

Corn ear-derived polynucleotide (cpd) #574.

Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023; structural gene; functional gene; regulatory gene;
corn ear-specific profile; gene transcription; gene expression;
hybrid plant; desirable trait expression; plant breeding program;
inheritance; desired characteristic; growth; development;
disease resistance; environmental adaptability; quality; yield;
multigene trait; plant; gene; ss.

Zea mays.

US6476212-B1.

05-NOV-2002.

14-MAY-1999; 99US-00313294.

26-MAY-1998; 98US-0086722P.

(INCY-) INCYTE GENOMICS INC.

Lalgudi RV, Ito LY, Sherman BK;

WPI; 2003-200840/20.

Novel purified corn-ear derived polynucleotide useful as hybridization probe for detecting polynucleotide in sample, and for identifying, evaluating, and altering desired characteristics associated with growth, development.

Example; SEQ ID NO 574; 390pp; English.

The present invention relates to the isolation of corn ear-derived polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022 and SATMON023. Some of the cdps uniquely identify structural, functional, and regulatory genes of corn ear. The polynucleotides sequences are useful for detecting cdps in a sample, for producing a corn ear-specific profile of gene transcription, for detecting altered gene expression in inbred or hybrid plants, and for screening several molecules for specific binding to the polynucleotide. The cdps are useful to identify, isolate, or extend identical or related corn-ear nucleic acid sequences from DNA libraries, and in nucleic acid amplification or hybridisation techniques to follow the expression of desirable traits through plant breeding programs. Preferably, the cdps are used to identify, evaluate, alter, or follow the inheritance of desired characteristics associated with growth and development, disease resistance, environmental adaptability, quality, and yield of corn. The cdps are also useful as molecular markers for studying inheritance and multigene traits in a plant breeding program. The cdps are useful for producing purified corn-ear polypeptides by recombinant techniques. They are also useful in diagnostic assays to detect or confirm conditions or diseases associated with abnormal levels of cdp expression. AX81541-AX89140 represent corn ear-derived polynucleotides (cdps) of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/psipdsIDEntry.html

SQ Sequence 270 BP; 85 A; 51 C; 70 G; 64 T; 0 U; 0 Other;

Query Match 13.7%; Score 199.6; DB 10; Length 270;

Best Local Similarity 83.7%; Pred. No. 4e-50;

Matches 226; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 92 GTGCTGACTGCAAAACAAAGGTCGAAGATGGGCTAGTGTAAATTTAGGTATCTTTATCT 151

Db 1 GCGCAGACTGCAAGTCAAGGTCCTCGATGGCAAGTGTGAATCTAGGTATCTTTATAT 60

QY 152 GCATGCATATGTTCTGGGATTCACAGGAGTCTCGGGGTACACATATCGAAGGTTGATCTG 211

Db 61 GCATGCATATGTTCTGGGATTCATAGAGCCCTGGGGTGCACATATCTAAGGTAAGATCTG 120

QY 212 CCATCTGGACACATGCTCCCGAGCAGGTTGCATTTATACAGTCAATGGGAATGATA 271

Db 121 CCATCTGGATACATGCTGCCAGAGCAAGTTGATTTTCAATCAATGGGAACGAA 180

QY 272 AAGCAATAGTTACTGGGAAGCAGAGTACCCCAAACTATGATAGAGTTGGAATTGAGA 331

Db 181 AAGCAATAGTTACTGGGAAGCAGAGTCCCTCTACTACGATAGGTTGGAATGAGA 240

QY 332 ATTTTATACGTGCAAGTATGAAGAGAGA 361

Db 241 ATTTTCATCCGTGCAAAATATAGGACAAGA 270

RESULT 13

AAC56661

ID AAC56661 standard; DNA; 408 BP.

XX AAC56661;

25-JAN-2001 (first entry)

Eucalyptus grandis transcription factor DNA sequence #532.

Plant; transcription factor; gene expression; eucalyptus; pine; acacia; poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
homeodomain zipper; LIM domain; AP2; ERBBs; zinc finger domain;
type 2 Cys2His2; CCAAT box element; MYB; ss.

Eucalyptus grandis.

WO2000053724-A2.

14-SEP-2000.

09-MAR-2000; 2000WO-US0006112.

11-MAR-1999; 99US-00266513.

18-AUG-1999; 99US-0149485P.

(GENE-) GENESIS RES & DEV CORP LTD.

(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Wood M, McGrath A, Shenk MA, Glenn M;

WPI; 2000-579369/54.

New isolated polynucleotide encoding a plant transcription factor for producing a plant e.g. a woody plant, preferably eucalyptus or pine, having modified gene expression or modified activity of a polypeptide.

Claim 1; Page 486; 747pp; English.

The present invention relates to novel plant transcription factors from Eucalyptus grandis or Pinus radiata. The present sequence is the coding sequence for one such transcription factor. The transcription factor may be used to produce a plant having modified gene expression such as a woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or to modify the activity of a polypeptide in a plant.

CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bZIP, bZIP family of G-box
CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, CCAAT box elements and
CC CC and CREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements and
CC MYB
XX
SQ Sequence 408 BP; 105 A; 96 C; 109 G; 98 T; 0 U; 0 Other;

Query Match 12.6%; Score 182.4; DB 3; Length 408;
Best Local Similarity 85.0%; Pred. No. 9.1e-45;
Matches 204; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 ATGACGAGAAAGCCACGCTCTTAAGGAGCTTAATCCCGGCATAGAAAGATTCTTGAA 60
Db 169 ATGACGAGAAAGCCACGCTCTTAAGGAGCTTAATCCCGGCATAGAAAGATTCTTGAA 228

QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAGA 120
Db 229 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAGA 288

QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCAATGTTCTGGGATTACAGAGT 180
Db 289 TGGGCTAGTGTCAATTTAGGATATTTATATGATGCAATGTTCTGGGATTACAGAGT 348

QY 181 CTGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCAG 240
Db 349 CTGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCAG 408

RESULT 14
ABL73349
ID ABL73349 standard; cDNA; 290 BP.
XX AC ABL73349;
XX
DT 14-MAY-2002 (first entry)
XX
DE Corn tassal-derived polynucleotide (cdps) SEQ ID NO:2723.
XX
XX Corn; corn tassal-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassal; gene; ss.
XX
XX Zea mays.
XX
XX OS
XX US2001051335-A1.
XX
XX PN
XX 13-DEC-2001.
XX
XX PD
XX PF 16-APR-1999; 99US-00294093.
XX
XX PR 21-APR-1998; 98US-0082567P.
XX
XX (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
XX Lalgudi RV, Ito LY, Sherman BK;
PI
XX WPI; 2002-163647/21.
DR
XX
XX Novel purified corn tassal-derived polynucleotide useful for determining
PT altered gene expression, to recover regulatory elements and to follow
PT inheritance of desirable characteristics through hybrid breeding
PT programs.
XX
XX Claim 1; SEQ ID NO 2723; 201pp; English.
XX
XX The present sequence describes a purified corn tassal-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences

CC encode corn tassal-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful in
CC the evaluation, and alteration of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC quality and yield, and as molecular markers for studying inheritance of
CC multigene traits in a plant breeding program. (I) can be used to produce
CC a tassal-specific profile of gene transcription, a transcript image, to
CC clone regulatory elements for use in transformation vectors, to express a
CC polypeptide, to identify, isolate or extend identical or related corn
CC tassal nucleic acid sequences from DNA libraries, in nucleic acid
CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences
XX
SQ Sequence 290 BP; 83 A; 63 C; 66 G; 55 T; 0 U; 23 Other;

Query Match 8.8%; Score 127.2; DB 6; Length 290;
Best Local Similarity 75.9%; Pred. No. 5.7e-28;
Matches 189; Conservative 0; Mismatches 57; Indels 3; Gaps 3;

QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAGA 120
Db 2 GGTCTTCTACGGCATCTCGAAGATAGAGAAATGGCAGACTGCAAGTCAAGGGTCTCTCGA 61

QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCAATGTTCTGG-GATTACAGGAG 179
Db 62 TGGGCAAGTGTGAATCTAGGTATCTTTATATGATGCAATGTTCTGGACATTCATAGAAG 121

QY 180 TCTCGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCA 239
Db 122 CTGGGGGTGACATATCTAAGGTAGATCTGCCAC-CCGGATACATGCTGCCAGCA 180

QY 240 GGTTCATTTATACAGTCAATCGGAAATGATAAAGCAATAGTTACTTGGGAAGCAGAGCT 299
Db 181 AGTTGCATTTATTCAATC-ANGGGACCCGAAANAGCAATCGNTATNGGNNACANACC 239

QY 300 ACCCCCAAA 308
Db 240 CCACCNACA 248

RESULT 15
ABQ98937
ID ABQ98937 standard; DNA; 587 BP.
XX AC ABQ98937;
XX
XX DT 04-NOV-2002 (first entry)
XX
XX DE Human ORF744 coding sequence.
XX
XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW Antiinflammatory; gene therapy; human; ORF; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
XX OS Homo sapiens.
XX
XX PN US2002082206-A1.
XX
XX PD 27-JUN-2002.
XX
XX PF 30-MAY-2001; 2001US-00867550.
XX
XX PR 30-MAY-2000; 2000US-0208427P.
XX
XX (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA

PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
FI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI; 2002-626554/67.
DR P-PSDB; ABP64374.
XX
XX New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
XX Claim 2; SEQ ID NO 1487; 78pp; English.
XX
XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP64681 and ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ Sequence 587 BP; 166 A; 151 C; 138 G; 130 T; 0 U; 2 Other;

Query Match 7.4%; Score 107.2; DB 6; Length 587;
Best Local Similarity 60.1%; Pred. No. 1.1e-21;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 23 CTAAGAGCTTAATGCCGCCATAGAAAGATTCTTGAAGGGCTTCTTAAACATCCAGAGA 82
DB 119 CTCAGAAGCTGAACGAGCAGCAGCAGCTCATCTCTCAAGCTTCTGAGGAGGAGGACA 178
QY 83 ACAGAGAAATGCTGACTGCTGCAAAACAAAAGGTCCAGATGGGCTAGTGTAAATTTAGGTA 142
DB 179 ACAAGTACTGCGCCGACTGCGAGGCCAAAGGTCTCTCGATGGGCTTCTTGAATATTTGGTG 238
QY 143 TCTTTATCTGCATGCAATGTTCTGGGATTTCACAGGAGTCTCGGGGTACACATATCGAAGG 202
DB 239 TGTATTATTTGCATCAGATGTCTGGAAATTCATAGAAATCTTGGGTTTATATATCCAGGG 298
QY 203 TTGATCTGCCACTCTGGACACATGCTCCCGAGCAGGTTGCATTTATACAGTCAATGG 262
DB 299 TCAATCAGTCAACCTAGACCAATGACAGCAGACAGATACAGTGCATGCAAGATATGG 358
QY 263 GAATCATTAAGCAATAGTTACTTGGGAGCAGAGCTACCCCAACTATGATAGA 318
DB 359 GAAATATAAGCAAGCACTACTCTATGAGCCAACTCTCCAGAGAACTTTGGAAGA 414

Search completed: September 27, 2005, 03:19:28

Job time : 785 secs


```
Db 181 CTCGGGTACACATATCGAAGTTGATCTGCCACTCTGGACACATGCTCCCGAGCAG 240
Qy 241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAATAGTACTTGGGAACGAGCTA 300
Db 241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAATAGTACTTGGGAACGAGCTA 300
Qy 301 CCCCCAACTATGATAGAGTGGAAATGAGAAATTTATACGTGCAAAATGATGAAGAGAG 360
Db 301 CCCCCAACTATGATAGAGTGGAAATGAGAAATTTATACGTGCAAAATGATGAAGAGAG 360
Qy 361 AGATGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGAGTGGAGAGAAACGGCGG 420
Db 361 AGATGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGAGTGGAGAGAAACGGCGG 420
Qy 421 AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATAGCATAGTACTAGTACTCTTAAT 480
Db 421 AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATAGCATAGTACTAGTACTCTTAAT 480
Qy 481 TTGTTTGAGGAGAGAAACTATTCCAGCATCTAGAAACAAAGAAATAATGTTGCTCAACG 540
Db 481 TTGTTTGAGGAGAGAAACTATTCCAGCATCTAGAAACAAAGAAATAATGTTGCTCAACG 540
Qy 541 AGAATAAATCTTCCGTGCTCCCAAGGACCCAGTCCAGTTCAGTTCAGTTCAGTTCAGT 600
Db 541 AGAATAAATCTTCCGTGCTCCCAAGGACCCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 600
Qy 601 ATGGAGTCTGAGCTACTCCAGTAGAGAGGAGAAACAGCAGTAATGTTGCAACGCA 660
Db 601 ATGGAGTCTGAGCTACTCCAGTAGAGAGGAGAAACAGCAGTAATGTTGCAACGCA 660
Qy 661 TCAGATCTCTCCAAAGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
Db 661 TCAGATCTCTCCAAAGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
Qy 721 TCGACTACAAATPACTCAGAGCAACTCTGCGCATCTCTGCGCATGATAACTCATGG 780
Db 721 TCGACTACAAATPACTCAGAGCAACTCTGCGCATCTCTGCGCATGATAACTCATGG 780
Qy 781 GCTGCTTTTCACTGCTGGAAGTGGTCAACCGGAGAGAAATGTCACGCCAAGCCT 840
Db 781 GCTGCTTTTCACTGCTGGAAGTGGTCAACCGGAGAGAAATGTCACGCCAAGCCT 840
Qy 841 GCTGAGAGCAGTTCCTCCAGCTTCATCTTCTGATCTTGGAGTTCGTTTAAAGACACA 900
Db 841 GCTGAGAGCAGTTCCTCCAGCTTCATCTTCTGATCTTGGAGTTCGTTTAAAGACACA 900
Qy 901 CCTAATTTAACTCAACAGACACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
Db 901 CCTAATTTAACTCAACAGACACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
Qy 961 GAGAGAGCAATATAGTATCGCTTTTGCATGTCATCAGACAGTTCGTTATGCTCGCT 1020
Db 961 GAGAGAGCAATATAGTATCGCTTTTGCATGTCATCAGACAGTTCGTTATGCTCGCT 1020
Qy 1021 CAGCAGCAAGCCCTTTTACATGCTGCGAGCAAGCTGCTGAGGCACTCCAAACGGCGTG 1080
Db 1021 CAGCAGCAAGCCCTTTTACATGCTGCGAGCAAGCTGCTGAGGCACTCCAAACGGCGTG 1080
Qy 1081 AATCAACAGCTATTGCTTAATGCTCTTAACTGCTTTCTGCAAAATGTTGTCAAACCCCGC 1140
Db 1081 AATCAACAGCTATTGCTTAATGCTCTTAACTGCTTTCTGCAAAATGTTGTCAAACCCCGC 1140
Qy 1141 GGTACCAAGTCCCGGATGATCAACCCGTAGTGGTCAAGCTGATCTCCAGAACTT 1200
Db 1141 GGTACCAAGTCCCGGATGATCAACCCGTAGTGGTCAAGCTGATCTCCAGAACTT 1200
Qy 1201 ATGCAAAACATCAATATGAACCAACATGAACAGAGCCCGCAACCGCAGAGAAC 1260
Db 1201 ATGCAAAACATCAATATGAACCAACATGAACAGAGCCCGCAACCGCAGAGAAC 1260
Qy 1261 ACTCTACAATACCCATCATCCAGTTTCTACAAATGGGTCAAGCTTAATCAAGTGAACGGT 1320
Db 1261 ACTCTACAATACCCATCATCCAGTTTCTACAAATGGGTCAAGCTTAATCAAGTGAACGGT 1320
```

```
Qy 1321 ATGACCCCAACTCAACGGTAAACCTCAGTCATCATCCGCAACCCCAACCAAGCACC 1380
Db 1321 ATGACCCCAACTCAACGGTAAACCTCAGTCATCATCCGCAACCCCAACCAAGCACC 1380
Qy 1381 ACACCATCTTCCAAATCAGGCAAGACTTTGATTTCTCTCTTGTGATGATGGAATGTC 1440
Db 1381 ACACCATCTTCCAAATCAGGCAAGACTTTGATTTCTCTCTTGTGATGATGGAATGTC 1440
Qy 1441 ACAAAACATTGA 1452
Db 1441 ACAAAACATTGA 1452

RESULT 2
BT000287 1612 bp mRNA linear PLN 19-SEP-2002
LOCUS Arabidopsis thaliana unknown protein (At5g54310) mRNA, complete cds.
DEFINITION
ACCESSION BT000287 GI:23198157
VERSION BT000287.1
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
1 (bases 1 to 1612)
REFERENCE
AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Sacou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (19-SEP-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Sacou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEN (SSP) Consortium members constructed and
sequenced the pENTR (ORF) clones using the RAFL cDNAs: Nguyen,M.,
Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H.,
Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W.,
Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Toroumi,M.,
Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S.,
Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
Location/Qualifiers
source 1..1612
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U25083"
/ecotype="Columbia"
/notes="This clone is in pENTR/SD-dTopo This is a cloned
PCR product using RIKEN clone RAFL09-80-M13 (AY099716) as
a template"
gene 1..1612
/gene="At5g54310"
CDS 1..1452
/gene="At5g54310"
/codon_start=1
```

```

/product="unknown protein"
/protein_id="AA015406.1"
/db_xref="GI:23198158"
/translation="MNEKANVSKELNARHKILEGLLKHPEENRECDCKTKGPRNASV
NLGIFCMQSGIHRSLGVIHSKVRGATLDTLWLEQVAFIQSMGNDKANSYWEALPP
NYDRVGIENFIRAKYERKEKVRSGEARSPPRVEQERKSVSGYEGHSHSPPVN
LPERERTIPASRTNNVAATRLNLPVPOGSOVIKPOOKMESAAATPVEREKOAVNVA
PASDPKVPFADLFNMLSMDSSTNTSATPGDTPADNWSAGFOSGSGTAEKIV
TAKPSSSPASSPDELFKDTNLTQQAPKDKVGDIMSLFEKTNLVSPPAMHQ
QVAMLAQOALYNAAKAAGGTGPNVQALNALNANWSNPNPGYQIPGHTNPGV
GQADLQKMNMMNMMNTRPAQPOENTLOYPSSFYTMGQANOVNGMTNPKTPQ
SSSATOPTSTPSSQSGKDFDPSLMDGMFKH"

```

ORIGIN

```

Query Match      100.0%; Score 1452; DB 8; Length 1612;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAAGCAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
DB 1 ATGACGAGAAAGCAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60

QY 61 GGGCTTTCTTAAACATCCAGAGAAACAGAGATGTGCTGACTGCAAAAACAAAGGTCCAAGA 120
DB 61 GGGCTTTCTTAAACATCCAGAGAAACAGAGATGTGCTGACTGCAAAAACAAAGGTCCAAGA 120

QY 121 TGGGCTAGTGTAAATTAGGTATCTTTATCTGATGCAATGTCTTGGGATTCACAGGAGT 180
DB 121 TGGGCTAGTGTAAATTAGGTATCTTTATCTGATGCAATGTCTTGGGATTCACAGGAGT 180

QY 181 CTCGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCAG 240
DB 181 CTCGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCAG 240

QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
DB 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300

QY 301 CCCCCAAACTATGATAGAGTTGAATTTGAGAAATTTATACGTCGCAAGATATGAGAGAG 360
DB 301 CCCCCAAACTATGATAGAGTTGAATTTGAGAAATTTATACGTCGCAAGATATGAGAGAG 360

QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTAGAGTCGAGCAGGAAAGCGCG 420
DB 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTAGAGTCGAGCAGGAAAGCGCG 420

QY 421 AATCTGTGAGAGAAAGTGGCGGGATATGAGCATGACATAGTAGTAGTCTGTGTAAT 480
DB 421 AATCTGTGAGAGAAAGTGGCGGGATATGAGCATGACATAGTAGTAGTCTGTGTAAT 480

QY 481 TTGTTTGAGGAGAGGAAAACCTATTCAGCGATCTAGAACAAAGAAATATGTTGCTGCAACG 540
DB 481 TTGTTTGAGGAGAGGAAAACCTATTCAGCGATCTAGAACAAAGAAATATGTTGCTGCAACG 540

QY 541 AGATAAATCTCCCGTCTCCCGAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
DB 541 AGATAAATCTCCCGTCTCCCGAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600

QY 601 ATGGAGTCTCAGCTACTCCAGTAGAGAGGAGGAGAAACAGCAGTAATGTTGCAACGCA 660
DB 601 ATGGAGTCTCAGCTACTCCAGTAGAGAGGAGGAGGAGAAACAGCAGTAATGTTGCAACGCA 660

QY 661 TCAGATCTCTCAAAGTGGATTTTGCTATCTGATCTGTTTAAACATGCTATCAATGGATGAT 720
DB 661 TCAGATCTCTCAAAGTGGATTTTGCTATCTGATCTGTTTAAACATGCTATCAATGGATGAT 720

QY 721 TCGACTACAAATACCTCAGAGGCAACTCTCGGGGATACTCTCCGATGATTAATCTCATGG 780
DB 721 TCGACTACAAATACCTCAGAGGCAACTCTCGGGGATACTCTCCGATGATTAATCTCATGG 780

QY 781 GCTGGCTTTTCAGTCTGCTGGAAGTGGTCAACCGGACAGAGAAATTTGTACAGCCCAAGCCT 840
DB 781 GCTGGCTTTTCAGTCTGCTGGAAGTGGTCAACCGGACAGAGAAATTTGTACAGCCCAAGCCT 840

```

```

QY 841 GCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGACTTTGAGATTGTTTAAAGACACA 900
DB 841 GCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGACTTTGAGATTGTTTAAAGACACA 900

QY 901 CCTAATTTAAACAACTCAAACAGCACAAAGATGTGAAAGGCGATATCATGAGCCTGTTT 960
DB 901 CCTAATTTAAACAACTCAAACAGCACAAAGATGTGAAAGGCGATATCATGAGCCTGTTT 960

QY 961 GAGAGACGATATAGTATCGCTTTTCCATGTCATCAGACACAGTTCTGCTATGCTCGCT 1020
DB 961 GAGAGACGATATAGTATCGCTTTTCCATGTCATCAGACACAGTTCTGCTATGCTCGCT 1020

QY 1021 CAGCAGCAAGCCCTTTTACATGGCTCGAGCAAGCTGCTGGAGGCACTCCAAACCGCGTG 1080
DB 1021 CAGCAGCAAGCCCTTTTACATGGCTCGAGCAAGCTGCTGGAGGCACTCCAAACCGCGTG 1080

QY 1081 AATCAACAAAGCTATTGCTTAATGCTCTTAAACGTAAGTCTTGAAGTGTGCAAAATGGTCAAAACCCCGC 1140
DB 1081 AATCAACAAAGCTATTGCTTAATGCTCTTAAACGTAAGTCTTGAAGTGTGCAAAATGGTCAAAACCCCGC 1140

QY 1141 GGCTACCAAGTCCCGGAAATGACTTAACCCGCTAGTGGTCAAGCTGATCTCAGAAACTT 1200
DB 1141 GGCTACCAAGTCCCGGAAATGACTTAACCCGCTAGTGGTCAAGCTGATCTCAGAAACTT 1200

QY 1201 ATGCAAAAACATGATATGAACGCAAAACATGAACAGAGACCCGCAACCCGCAAGAGAAC 1260
DB 1201 ATGCAAAAACATGATATGAACGCAAAACATGAACAGAGACCCGCAACCCGCAAGAGAAC 1260

QY 1261 ACTCTACAAATCCCATCATCCAGTTTCTTACAACATGGGTCAAGCTAATCAAGTGAACGGT 1320
DB 1261 ACTCTACAAATCCCATCATCCAGTTTCTTACAACATGGGTCAAGCTAATCAAGTGAACGGT 1320

QY 1321 ATGACCCCAACTCAACCGGTAACCTCAGTCATCATCCGCAACCCCAACCAACAGCACC 1380
DB 1321 ATGACCCCAACTCAACCGGTAACCTCAGTCATCATCCGCAACCCCAACCAACAGCACC 1380

QY 1381 ACACCATCTTCAACATCAGGCAAGACTTTGATTTCTTCTTCTGATGATGGAATGTTTC 1440
DB 1381 ACACCATCTTCAACATCAGGCAAGACTTTGATTTCTTCTTCTGATGATGGAATGTTTC 1440

QY 1441 ACAAAACATTGA 1452
DB 1441 ACAAAACATTGA 1452

```

```

RESULT 3
AY099716 LOCUS
DEFINITION Arabidopsis thaliana unknown protein (At5g54310) mRNA, complete
cds
ACCESSION AY099716
VERSION AY099716.1 GI:20466499
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1777)
AUTHORS Nguyen.M., Karlin-Neumann.G., Southwick.A., Lam.B., Miranda.M.,
Palm.C.J., Bowser.L., Jones.T., Banh.J., Carninci.P., Chen.H.,
Cheuk.R., Chung.M.K., Hayaishizaki.Y., Ishida.J., Kamiya.A.,
Kawai.J., Kim.C., Lin.J., Liu.S.X., Narusaka.M., Pham.P.K.,
Sakano.H., Sakurai.T., Satou.M., Seki.M., Shinn.P., Yamada.K.,
Shinozaki.K., Ecker.J., Theologis.A. and Davis.R.W.
Direct Submission
Submitted (24-APR-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu
RIKEN Genomic Sciences Center (GSC) members carried out the

```

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIVEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carrinci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J., Bower, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J. Theologian, A. and Davis, R. W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

[illegible]

Location/Qualifiers
1. .1777

```

I. : I / / /
/organism="Arabidopsis thaliana"

```

```

I. : I / / /
/organism="Arabidopsis thaliana"

```

```
/mol_type="mrna"
```

```
/db xref="taxon:3702"
```

```
/chromosome="5"
```

```
/clone="RAFL09-80-M13 (R25083)"
```

/ecotype="Columbia"

/note="This clone is in pBluescript"

1. 1777

/gene="At5g54310"

151. .160

/gene="At5g54310"

```
/codon_start=1
```

/product="unknown protein"

```
/protein_id="AAM20567.1"
```

```
/db_xref="GI:20466500"
```

/translacion="mwekaskelmarukilteggulapfenlecaudckinfpwas
 nlgi ft cmqcsghrslyvhi skvatsadltwlpqvaf t osmgndkxanswtaepal
 lyferertien irakyeeknwvgerkarspvrveqcsyvesgvpgechsspnv
 lfyeerkti pasrtrenvvaatrl nlpvpqpsvotlkpookmesatpvarekoavna
 pasdppkvpatdl fwnlsmdststntseatpgdt paddnswagcfsgsqtaekiv
 tapassppspasfedeu fkdtpnl tltlqqapkgdgt mslcfepv fspamhoo
 qvamlaoqqal ymaakaaqzsqvngvqaoatana lnavas nspgvyi lpgvntpv
 goad lqkqmnwmmamtrpaqent lqf vpsssfytmqaoqvngmtpnsgtkp
 ssasatpstsstsssgskpdpsssi mdgwtfw"

ORIGIN

Query Match	100.0%;	Score 1452;	DB 8;	Length 1777;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1452;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAACGAGAAAGCCAAAGCTCTCTTAAGGAGCTTAATGCCGCCCATGAAAGATTCCTGAA	60	
Db	151	ATGAACGAGAAAGCCAAAGCTCTCTTAAGGAGCTTAATGCCGCCCATGAAAGATTCCTGAA	210	
QY	61	GGGCTTCTTAAACATCCAGAGAAACAGAGAATGTGCTGACTGCAGAAACAAAGGTCCTCAAGA	120	
Db	211	GGGCTTCTTAAACATCCAGAGAAACAGAGAATGTGCTGACTGCAGAAACAAAGGTCCTCAAGA	270	
QY	121	TGGCTAGTGTTAATTTAGTATCTTTATCTGCATGCAATGTTCTGGATTCACAGAGT	180	
Db	271	TGGCTAGTGTTAATTTAGTATCTTTATCTGCATGCAATGTTCTGGATTCACAGAGT	330	
QY	181	CTGGGGTACACATATCGAAGTTGATCTGCCACTCTGGACACATGGCTCCCGGACG	240	
Db	331	CTGGGGTACACATATCGAAGTTGATCTGCCACTCTGGACACATGGCTCCCGGACG	390	
QY	241	GTTCGATTTATACAGTCAATGGGAAATGATAAGCAATAGTTACTGGGAAGCAGAGCTA	300	
Db	391	GTTCGATTTATACAGTCAATGGGAAATGATAAGCAATAGTTACTGGGAAGCAGAGCTA	450	
QY	301	CCCCAAACTATGATAGAGTTGGAAATTTGAGAAATTTTATAGCTGCAAAAGTATGAAGAGAAG	360	
Db	451	CCCCAAACTATGATAGAGTTGGAAATTTGAGAAATTTTATAGCTGCAAAAGTATGAAGAGAAG	510	
QY	361	AGATGGGTTTTCTAGAGGGGAAAAGGCTTAGATCACTCTTAGAGTCGACAGGAACGGCGG	420	

```

RESULT 4
AY088913 LOCUS 1794 bp mRNA linear PLN 14-APR-2003
Arabidopsis thaliana clone 99519 mRNA, complete sequence.
AY088913 DEFINITION
AY088913 ACCESSION
AY088913.1 GI:21407687
FLI CDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1794)
Hae, B.J., Volfovsky, N., Town, C.D., Troughan, M., Alexandrov, N.,
Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
22088475
12093376
2 (bases 1 to 1794)
Brover, V., Troughan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1794)
Brover, V., Troughan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessments of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
Location/Qualifiers
1..1794
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="99519"
167..1618
/codon_start=1
/product="ARF GAP-like zinc finger-containing protein
ZIG3"
/db_xref="GI:21618169"
/protein_id="AAM67219.1"
/db_xref="M5EKANVSKELNARHKLEGLKHPENRECAADCKTGPWASV
NIGIFCMQCSGHRSLGVHISKVRSATLDLWLPQVAFIQSGMNDKANSYWEAELPP
NYDRVGIEFIKAKYBEKRWKSGEKARSPRVEQRRKSVTSFGYEHGSSSPVN
LFEERTIIPATRNVAATRLNLPVPPQSQVVKPQQKMSAATPVREKQAVNVA
PASDPKVPFADTLFNLGMDSTTNTSEATPGDTPADNSWAGSQSGQTAEKIV
TAKPASSSPPASSDFDLFKOTPNLTQQAPKDKVGDIMSLIFEXTNIVSPFAMHQ
QVAMLAQQALYMAAKAGGTGNGVNOQAIALNANVANSNPGYOIPGMWPNVG
QADLQKLNQNMNMMNTRPAQPDENTLOVPSSSFYTMGQANQVNGWTPNSTGKP
SSSATQPTSTPSSQSGKQDFDSSLMDGMFTKH"

FEATURES
source
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="99519"
167..1618
/codon_start=1
/product="ARF GAP-like zinc finger-containing protein
ZIG3"
/db_xref="GI:21618169"
/protein_id="AAM67219.1"
/db_xref="M5EKANVSKELNARHKLEGLKHPENRECAADCKTGPWASV
NIGIFCMQCSGHRSLGVHISKVRSATLDLWLPQVAFIQSGMNDKANSYWEAELPP
NYDRVGIEFIKAKYBEKRWKSGEKARSPRVEQRRKSVTSFGYEHGSSSPVN
LFEERTIIPATRNVAATRLNLPVPPQSQVVKPQQKMSAATPVREKQAVNVA
PASDPKVPFADTLFNLGMDSTTNTSEATPGDTPADNSWAGSQSGQTAEKIV
TAKPASSSPPASSDFDLFKOTPNLTQQAPKDKVGDIMSLIFEXTNIVSPFAMHQ
QVAMLAQQALYMAAKAGGTGNGVNOQAIALNANVANSNPGYOIPGMWPNVG
QADLQKLNQNMNMMNTRPAQPDENTLOVPSSSFYTMGQANQVNGWTPNSTGKP
SSSATQPTSTPSSQSGKQDFDSSLMDGMFTKH"

CDS
167..1618
/codon_start=1
/product="ARF GAP-like zinc finger-containing protein
ZIG3"
/db_xref="GI:21618169"
/protein_id="AAM67219.1"
/db_xref="M5EKANVSKELNARHKLEGLKHPENRECAADCKTGPWASV
NIGIFCMQCSGHRSLGVHISKVRSATLDLWLPQVAFIQSGMNDKANSYWEAELPP
NYDRVGIEFIKAKYBEKRWKSGEKARSPRVEQRRKSVTSFGYEHGSSSPVN
LFEERTIIPATRNVAATRLNLPVPPQSQVVKPQQKMSAATPVREKQAVNVA
PASDPKVPFADTLFNLGMDSTTNTSEATPGDTPADNSWAGSQSGQTAEKIV
TAKPASSSPPASSDFDLFKOTPNLTQQAPKDKVGDIMSLIFEXTNIVSPFAMHQ
QVAMLAQQALYMAAKAGGTGNGVNOQAIALNANVANSNPGYOIPGMWPNVG
QADLQKLNQNMNMMNTRPAQPDENTLOVPSSSFYTMGQANQVNGWTPNSTGKP
SSSATQPTSTPSSQSGKQDFDSSLMDGMFTKH"

Query Match 99.7%; Score 1447.2; DB 8; Length 1794;

```

```

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1449; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGACGAGAGGCAACGCTCTTAAGGAGCTTAAATGCCGCATAGAAAGATTCTTGAA 60
DB 167 ATGACGAGAGGCAACGCTCTTAAGGAGCTTAAATGCCGCATAGAAAGATTCTTGAA 226
QY 61 GGGCTTCTTAAACATCCAGAGACAGAGAAATGTGTGCTGCAATGCTGCAAAACAAAGGTCCAAGA 120
DB 227 GGGCTTCTTAAACATCCAGAGACAGAGAAATGTGTGCTGCAATGCTGCAAAACAAAGGTCCAAGA 286
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGCT 180
DB 287 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGCT 346
QY 181 CTCGGGGTACACATATCCAAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGACGAG 240
DB 347 CTCGGGGTACACATATCCAAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGACGAG 406
QY 241 GTTGCAATTTATACAGTCAATCGGAAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB 407 GTTGCAATTTATACAGTCAATCGGAAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA 466
QY 301 CCCCAAACTATGATAGAGTTGGAAATTCAGAAATTTTATACGTGCAAACTATCAAGAGAG 360
DB 467 CCCCAAACTATGATAGAGTTGGAAATTCAGAAATTTTATACGTGCAAACTATCAAGAGAG 526
QY 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAAACGGCGG 420
DB 527 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAAACGGCGG 586
QY 421 AAATCTGTGGAGAGAAAGTGGGCGCGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 480
DB 587 AAATCTGTGGAGAGAAAGTGGGCGCGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 646
QY 481 TTGTTTGGAGAGAGAAACCTATTTCACGATCTAGAACAGAAATTAATGTTCTGCAACG 540
DB 647 TTGTTTGGAGAGAGAAACCTATTTCACGATCTAGAACAGAAATTAATGTTCTGCAACG 706
QY 541 AGAATAAATCTTCCCGTCTCCCAAGACCCAGTCTAGGTTTAAAGCCACAGCAGAGAA 600
DB 707 AGAATAAATCTTCCCGTCTCCCAAGACCCAGTCTAGGTTTAAAGCCACAGCAGAGAA 766
QY 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAACAGCAGTAATTTGCAACACGCA 660
DB 767 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAACAGCAGTAATTTGCAACACGCA 826
QY 661 TCAGATCTCTCAAAGGTGGATTTTCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
DB 827 TCAGATCTCTCAAAGGTGGATTTTCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 886
QY 721 TCGACTTACAAATACCTCAGAGGCAACTCTCTGGCGATATCTCTGCCGATGATAACTCATGG 780
DB 887 TCGACTTACAAATACCTCAGAGGCAACTCTCTGGCGATATCTCTGCCGATGATAACTCATGG 946
QY 781 GCTGGCTTTCACTGTCTGGAAGTGGTCAACCGCAGAGAAATTTGTCACAGCAGACCT 840
DB 947 GCTGGCTTTCACTGTCTGGAAGTGGTCAACCGCAGAGAAATTTGTCACAGCAGACCT 1006
QY 841 GCTGAGAGAGTTCCTCTCCAGCTTCACTTCTGACTTTGAGGATTTGTTTAAAGGACACA 900
DB 1007 GCTGAGAGAGTTCCTCTCCAGCTTCACTTCTGACTTTGAGGATTTGTTTAAAGGACACA 1066
QY 901 CCTAATTTAAACAACTCAACAGCACAAAGAGATGTGAAGGCGATATCATGAGCCTGTTT 960
DB 1067 CCTAATTTAAACAACTCAACAGCACAAAGAGATGTGAAGGCGATATCATGAGCCTGTTT 1126
QY 961 GAGAAAGCAATATAGTATCGCTTTTGGCATGCAATGCAACAGGTTGCTATGCTCGCT 1020
DB 1127 GAGAAAGCAATATAGTATCGCTTTTGGCATGCAATGCAACAGGTTGCTATGCTCGCT 1186
QY 1021 CAGCAGCAAGCCCTTTTACATGCTGAGGAAAGTCTGAGGCACTCCAAACGGCGTG 1080

```

```

Db      1187  CAGCAGCAAGCCCTTTACATGGCTGCAGCGAAGAGCTGCTGGAGGCACTCCAAAACGGCGTG 1246
QY      1081  AATCAACAGCTATTGCTTAATGCTCTTAAGCTAGCTTCTGCAAAATGGTCAAAACCCCGC 1140
Db      1247  AATCAACAGCTATTGCTTAATGCTCTTAAGCTAGCTTCTGCAAAATGGTCAAAACCCCGC 1306
QY      1141  GGCTACAGATCCCGGAATGACTAACCCCGTAGGTGGTCAAGCTGATCTCCAGAAACTT 1200
Db      1307  GGCTACAGATCCCGGAATGACTAACCCCGTAGGTGGTCAAGCTGATCTCCAGAAACTT 1366
QY      1201  ATGCAAAAATGATATGAAAGCAACATGACACGAGACCCGCAACCCGCAAGCAAGAAC 1260
Db      1367  ATGCAAAAATGATATGAAAGCAACATGACACGAGACCCGCAACCCGCAAGCAAGAAC 1426
QY      1261  ACTCTACAATACCCATCATCTCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAAGCT 1320
Db      1427  ACTCTACAATACCCATCATCTCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAAGCT 1486
QY      1321  ATGACCCCAAACTCAACCGGTAAACCTCAGTCAATCATCCGCAACCCCAACCAAGCACC 1380
Db      1487  ATGACCCCAAACTCAACCGGTAAACCTCAGTCAATCATCCGCAACCCCAACCAAGCACC 1546
QY      1381  ACACCATCTTCAATCAGCGCAAGACTTTGATTTCTTCTTCTGATGATGGAATGTTTC 1440
Db      1547  ACACCATCTTCAATCAGCGCAAGACTTTGATTTCTTCTTCTGATGATGGAATGTTTC 1606
QY      1441  ACAAACATTGA 1452
Db      1607  ACAAACATTGA 1618

RESULT 5
AF184144
LOCUS   Arabidopsis thaliana ARF GAP-like zinc finger-containing protein
DEFINITION
ACCESSION   AF184144
VERSION     AF184144.1
KEYWORDS    GI:10441351
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
REFERENCE   1 (bases 1 to 1492)
AUTHORS    Jensen,R.B., Lykke-Andersen,K., Frandsen,G.I., Nielsen,H.B.,
            Haseloff,J., Jespersen,H.M., Mundy,J. and Skriver,K.
TITLE      Promiscuous and specific phospholipid binding by domains in ZAC, a
            membrane-associated Arabidopsis protein with an ARF GAP zinc finger
            and a C2 domain
JOURNAL    Plant Mol. Biol. 44 (6), 799-814 (2000)
MEDLINE    21070549
PUBMED     11202441
REFERENCE   2 (bases 1 to 1492)
AUTHORS    Andersen,K.L., Jensen,R.B., Jespersen,H.M. and Skriver,K.
TITLE      Direct Submission
JOURNAL    Submitted (10-SEP-1999) Dept. of Protein Chemistry, Inst. of
            Molecular Biology, University of Copenhagen, Oe. Farimagsgade 2A,
            Copenhagen 1353, Denmark
FEATURES   Location/Qualifiers
            1..1492
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /cultivar="Columbia"
            /db_xref="taxon:3702"
            /chromosome="5"
            <1..1436
            /note="similar to protein encoded by GenBank Accession
            Number AB010693"
            /codon_start=3
            /product="ARF GAP-like zinc finger-containing protein
            ZIG3"
            /protein_id="AA017004.1"

CDS

```

```

/db_xref="GI:10441352"
/translation="VSKELNARHKILLEGLLKHPENRECAKCTKGRWASVNLGIFI
CMQSGIHRISGVHISKVRSATLDTLWPEQVAFIQSMGNDRKANSYWEALPENYDRVG
IENIRAKYEKRWVSGEKARSPPRVEQERKSVRSGPGVEHGHSSPNVLPERRK
TIPASRTNNVATRIINLPVPPQSPQVIFKQKMSASATPVERKQVNVNAPADPP
KVDPATDLFNLMSDDSTTNTSEATPGDTPADDNSWAGFSAGSGOTAEKIVTAKPAE
SSSPASSSDFEDLFKDTPLNLTQQAQKVDKIDMSLPFKTNI VSPFMHQOQVAMLA
QQQALYMAAKAAGGTGNGVNOQAIAANALVANASWNPGGYQI PGMTNPPVGGADLQ
KLMQNMNANMNTRPAPQPOENTLOYPSSPYTMGOANOVNGMTPNSTGKQSSSATO
PTSTTPSSQSKGDFDFSLMDMGMTKH"

ORIGIN
Query Match      98.9%; Score 1436; DB 8; Length 1492;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17  ACGTCTCTAAGAGCTTAATGCCCGCCCATAGAAAGATTTCTTGAAGGGCTCTTTAAACATC 76
Db      1  ACGTCTCTAAGAGCTTAATGCCCGCCCATAGAAAGATTTCTTGAAGGGCTCTTTAAACATC 60
QY      77  CAGAGAACAGAGAATGTGCTGACTGCAAAAACAAAAGGTCCAAAGATGGGCTAGTGTAAAT 136
Db      61  CAGAGAACAGAGAATGTGCTGACTGCAAAAACAAAAGGTCCAAAGATGGGCTAGTGTAAAT 120
QY      137  TAGGTATCTTTATTCGATGCAATGTTTCTGGGATTCACAGAGTCTCGGGGTACACATAT 196
Db      121  TAGGTATCTTTATTCGATGCAATGTTTCTGGGATTCACAGAGTCTCGGGGTACACATAT 180
QY      197  CGAAGTTTCGATTCGCACCTCTGGACACATGGCTCCCGGACGAGTTCGATTATACAGT 256
Db      181  CGAAGTTTCGATTCGCACCTCTGGACACATGGCTCCCGGACGAGTTCGATTATACAGT 240
QY      257  CAATGGGAATGATAAAGCAATAGTTACTCGGAAGCAGAGCTACCCCAAACTATGATA 316
Db      241  CAATGGGAATGATAAAGCAATAGTTACTCGGAAGCAGAGCTACCCCAAACTATGATA 300
QY      317  GAGTTGGAATTCAGAAATTTTATACGTGCAAAATAGTAGTCTGTAAATTTGTTGAGGAGAGA 376
Db      301  GAGTTGGAATTCAGAAATTTTATACGTGCAAAATAGTAGTCTGTAAATTTGTTGAGGAGAGA 360
QY      377  GGGAAAGGCTAGATCACTCTAGCTGAGAGGAGGACGGGAAATCTGTGGAGAGAA 436
Db      361  GGGAAAGGCTAGATCACTCTAGCTGAGAGGAGGACGGGAAATCTGTGGAGAGAA 420
QY      437  GTGGCCCGGATATGAGCATGAGATGATAGTCTGTAAATTTGTTGAGGAGAGA 496
Db      421  GTGGCCCGGATATGAGCATGAGATGATAGTCTGTAAATTTGTTGAGGAGAGA 480
QY      497  AAATCTATTCAGATCTAGAAACAAGAAATATGTTGTCGCAACGAGATTAATCTTCCG 556
Db      481  AAATCTATTCAGATCTAGAAACAAGAAATATGTTGTCGCAACGAGATTAATCTTCCG 540
QY      557  TGCCTCCCGAGGACCCAGTCCAGTGTATAAGCCACAGCAAAATGAGTCTGCAGCTA 616
Db      541  TGCCTCCCGAGGACCCAGTCCAGTGTATAAGCCACAGCAAAATGAGTCTGCAGCTA 600
QY      617  CTCAGTAGAGAGGAGAGAAACAAGCAGTAAATTTGTCACCAAGCATCAGATCTTCAAAG 676
Db      601  CTCAGTAGAGAGGAGAGAAACAAGCAGTAAATTTGTCACCAAGCATCAGATCTTCAAAG 660
QY      677  TGGATTTTGTCTACTGATCTGTTTAACTGCTATCAATGATGATTCGACTCAAAATACCT 736
Db      661  TGGATTTTGTCTACTGATCTGTTTAACTGCTATCAATGATGATTCGACTCAAAATACCT 720
QY      737  CAGAGGCAACTCTGCGGATCTCTGCGATGATTAATCTCATGGCTGGCTTTCAGTCTG 796
Db      721  CAGAGGCAACTCTCTGCGGATCTCTGCGATGATTAATCTCATGGCTGGCTTTCAGTCTG 780
QY      797  CTGGAAGTGGTCAAAACGGCAGAGAAAATTTGTACAGCCAAAGCTGCTGAGAGCAGTCTTC 856
Db      781  CTGGAAGTGGTCAAAACGGCAGAGAAAATTTGTACAGCCAAAGCTGCTGAGAGCAGTCTTC 840
QY      857  CTCAGGCTTCAATCTTCTGACTTTTGAGGATTTGTTTAAAGCACACACCTAATTTTAACTC 916

```

Db	841	CTCCAGCTTCATCTTCTGACTTGTGAGATTTGTTTAAAGGACACACCTAATTTAAACAATC	900
Qy	917	AACAAGCACAAGAGTGAAGGCGATATCATGAGCTGTTTGAAGAGCAAGATATAG	976
Db	901	AACAAGCACAAGAGTGAAGGCGATATCATGAGCTGTTTGAAGAGCAAGATATAG	960
Qy	977	TATCGCTTTTGCATGATCAGCAAGTTCGTATGCTCGCTCAGCAGCAGCCCTTT	1036
Db	961	TATCGCTTTTGCATGATCAGCAAGTTCGTATGCTCGCTCAGCAGCAGCCCTTT	1020
Qy	1037	ACATGCTGCGAGGAAAGCTGCTGAGGCACTCCAAACCGCGTGAATCAACAGCTATTG	1096
Db	1021	ACATGCTGCGAGGAAAGCTGCTGAGGCACTCCAAACCGCGTGAATCAACAGCTATTG	1080
Qy	1097	CTAATGCTTTAAGTATGCTTTCGAAATTTGTTCAACCCCGCGCTACAGATCCCG	1156
Db	1081	CTAATGCTTTAAGTATGCTTTCGAAATTTGTTCAACCCCGCGCTACAGATCCCG	1140
Qy	1157	GAATGACTAACCCCGTAGGTGCTCAAGCTGATCTCCAGAACTTATGCAAAACATGA	1216
Db	1141	GAATGACTAACCCCGTAGGTGCTCAAGCTGATCTCCAGAACTTATGCAAAACATGA	1200
Qy	1217	TGAACGCAACATGAAACACGAGACCCGCAACCCGCAAGAGAACACTTACATATCCCAT	1276
Db	1201	TGAACGCAACATGAAACACGAGACCCGCAACCCGCAAGAGAACACTTACATATCCCAT	1260
Qy	1277	CATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGTATGACCCCAACTCA	1336
Db	1261	CATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGTATGACCCCAACTCA	1320
Qy	1337	CCGTTAACTCTAGTATCATCGCAACCCCAACCAACGACCAACCATCTTCAAT	1396
Db	1321	CCGTTAACTCTAGTATCATCGCAACCCCAACCAACGACCAACCATCTTCAAT	1380
Qy	1397	CAGCAAGACTTTGATTTCTCTCTTCTGATGATGAATGTTTCAACAACATTTGA	1452
Db	1381	CAGCAAGACTTTGATTTCTCTCTTCTGATGATGAATGTTTCAACAACATTTGA	1436
RESULT 6			
AK067240			
LOCUS	AK067240	1451 bp	linear
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J013096N21, full insert sequence.		PLN 24-JUL-2003
ACCESSION	AK067240		
VERSION	AK067240.1		GI:32977258
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE			
AUTHORS			

JOURNAL MEDLINE PUBLISHED REFERENCE AUTHORS

Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2 (bases 1 to 1451)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imanura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurotsuki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
 Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28K full-length cDNA clones from japonica rice.
 URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
 PAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hiraoka,T., Hori,F., Iida,J., Imanura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
 Location/Qualifiers
 1. .1451
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J013096N21"

TITLE JOURNAL

COMMENT

FEATURES source

ORIGIN

Query Match 21.1%; Score 307; DB 8; Length 1451;
 Best Local Similarity 59.2%; Pred. No. 4.1e-74;
 Matches 614; Conservative 0; Mismatches 390; Indels 33; Gaps 4;
 1 ATGAACGAGAAAGCCAAACGCTCTCTTAGGAGCTTATATGCCCGCATGAAAGATTCTTGAA 60

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice


```

Db      278  ATGACGAGAGGCTCGCTCCACGAGGCTCAACCCAGCACAGAGATATTGGA 337
Qy      61  GGGCTTTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTTCACA 120
Db      338  GGCCTCTCGGGTGCCTGAGAAATAGAGAAATGTGCAGCTGCAAGTCAAAAGGTCCTCG 397
Qy      121  TGGGCTAGTGTAAATTTAGTGTCTTTATCTGCAATGCAATGTTCTGGGATTCACAG 180
Db      398  TGGCAGAGTGAATCTTGGGATCTTTATATGATGATGATGTTCTTGGAAATTCATAG 457
Qy      181  CTGGGGTACACATATATGAAAGTTTGCATCTGCCACTCTGGACACATGGCTCCCCG 240
Db      458  CTCGGGTACATATATCTAAGTAAGATCGCTACCTGATACATGGCTGCCAGACAA 517
Qy      241  GTTGCAATTTATACAGTCAATGCGGAATGATAAGCAAAATAGTTACTCGGAGACAG 300
Db      518  GTTGCAATTTATCAATCAATGCGGAATGAAAGTCAAAATAGTACTTGGGAAGCGG 577
Qy      301  CCCCCAACTATGATAGTTGGAATTTGAGAAATTTATACGTGCAAGATGTAAGA 360
Db      578  CCCCCTAATATATAGAGTTGGGATAGAGAACTTCATCCGTGCAAAATACGAGG 637
Qy      361  AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTTCGAGCAGGAA 420
Db      638  AGATGGATACCGAGGATGGGTCTCATCAGACCATCATCTGTGCTAGAGATGAG 697
Qy      421  AATCTGTGGAGAGAAAGTGGCGGGATATGAGATGGAATAGTAGTAGTCTCTGTA 480
Db      698  TCAGAGTCACAACTAGTTTAACAGGGGTGGCCATAATCAAAG-----ATCT 745
Qy      481  TTCTTTGAGAGAGGAAACTATTTCCAGCATCTAGACAGAAATATGTTGCTGCA 540
Db      746  TCATTTGAGCAACACCGCATCTTCCAGCTGCTGAGTGGAGCAAAATGCTCCAG 805
Qy      541  AGAATAAATCTTCCGTGCTCCCAAGGACCCAGTCAGGTATATAAGCCACACG 600
Db      806  AGGACGCCCAACCCAGGACCAACATCAGCCAAAGACCAACCATCAGTTCCTCA 865
Qy      601  ATGGAGTCTGAGTACTCAGTAGAGAGAGGAGAAACAGCA---GTAATGTTGCA 657
Db      866  CCTCCTCAACCCGAGAAATCACCTCTTAATGCAACACCAACCAAGGTTGAG 925
Qy      658  GCATCAGATCCTCAAGGTGGATTTTGTCTGATGATCTGTTTAAACATGCTATCA 717
Db      926  GTTGACCACTCTTAAGTTGACTATGCAACTGATCTTCAACATGTTATCATGG 985
Qy      718  GATTCCACTACAAATACCTCAGAGGCAACTCTCTGGCGATACCTCTCGCATGA 777
Db      986  GGAACAACCGAGAAAGAGCAGAGTCA-----TCTTCTAATGATGATGCA 1033
Qy      778  TGGGCTGGCTTTTCACTGCTGGAAGTGGTCAACCGCAGAGAAATTTGTCAGC 837
Db      1034  TGGGAGGCTTCCAGTCTGAGAGCCAGTACCTAGCTCAGACAAAAGGATTTCTG 1093
Qy      838  CTGCTGAGAGCAGTTCTCTCCAGCTTCACTTCTTGAC-----TTTGAGGAT 891
Db      1094  CCAGTAGAAGTAGGCCCACTTACATCAGGATAGAGACTTGTTTAAGATTCAC 1153
Qy      892  AAGGACACACTAATTTAACTCAACAGCACCAGAAAGATGTGAAGCCGATATCAT 951
Db      1154  GCTGTGACATGATCTCTCAGCTCCAGCTGCCCGCACAGTAATGTGAAGATGAT 1213
Qy      952  AGCCTGTTGAGAGAGAAATATAGTATCCCTTTTGGCCATGATCAGCAAGGTTG 1011
Db      1214  AGTCTGTTTGAAGTCCAGTATGTTTGGCCATATGCTGTCCAGCAGCAGCAACT 1273
Qy      1012  ATGCTCGCTCAGAGCA 1028
Db      1274  TTTATGACCCCCAGCA 1290

```

. RESULT 7

AK067125
LOCUS
DEFINITION
ORYZA SATIVA (JAPONICA CULTIVAR-GROUP) cDNA clone: J013098J09, full insert sequence.
AK067125
AK067125.1 GI:32977143
FLI CDNA; CAP trapper.
ORYZA SATIVA (JAPONICA CULTIVAR-GROUP)
ORYZA SATIVA (JAPONICA CULTIVAR-GROUP)
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team., Kikuchi, S., Satoh, K., Negata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 1741)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Kagawa, I., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kishimoto, N., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Negata, T., Niihara, J., Nishikawa, K., Nomura, M., Numasaki, R., Nakikawa, R., Niihara, J., Nishi, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Sugano, S., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Teggami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE
JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel.81-29-838-7007, Fax:81-29-838-7007)

COMMENT
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsumoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Mateubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, T., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akai, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
1. .1741
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013098J09"

FEATURES source

ORIGIN

Query Match 21.1%; Score 307; DB 8; Length 1741;
Best Local Similarity 59.2%; Pred. No. 4.1e-74;
Matches 614; Conservative 0; Mismatches 390; Indels 33; Gaps 4;

QY 1 ATGAACGAGAAAGCCAACTCTTAAGGAGCTTAATGCCCGCCATAGAAGATCTTCGAA 60
DB 43 ATGAACGAGAAAGCCCTCGCTCCAGGAGCTCAACGCCAAGCACAAGAGATATGGAA 102
QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGTGATCGTCAAAACAAAGAGTCCAA 120
DB 103 GGCCTCTCGCGCTGCCTGAGAATAGAGAATGTGCAGACTGCAAGTCAAGGGTCTCTGA 162
QY 121 TGGCTAGTGTAAATTTAGTATCTTTCTCATGCAATGTCTGGGATTCACAGGAGT 180
DB 163 TGGCAAGTGTAAATCTTGGGATCTTTATGATGTCAGTGTCTTGGAAATTCATAGAAGC 222
QY 181 CTCGGGTACACATATCGAAGTTCGATCTGCACCTCTGGACATGCTCCCGCAGCAG 240
DB 223 CTCGGGTACATATCTTAGTAAAGTTCGGTACCTGGATACATGCTGCCAGGCA 282
QY 241 GTTGCAATTTATACGTCAATGGAAATGATAAGCAATAGTTACTTGGGAAGCAGAGCTA 300
DB 283 GTTGCAATTTATCAATCAATGGGAATGAAAGTCAATAGCTACTTGGGAAGCGGAGCTG 342
QY 301 CCCCCAACTATGATAGTGTGAATTTGGAATTTTATACGTGCAAGTATGAAGAGAG 360
DB 343 CCCCCAACTATGATAGTGTGGATAGAGAACTTCATCCCGTCAAAATACGAGGACAA 402
QY 361 AGATGGTGTTCAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCCAGCAGGAACGGCGG 420
DB 403 AGATGGATACCGAGAAATGGTTCATCAGACCATCATCTGTGTCTAGAGATGAGAGAGC 462
QY 421 AAATCTGTGGAGAAAGTGGCGGGATATAGCATGGAATAGTATGATGCTCTTAAT 480
DB 463 TCAGAGTCACAACTAGTGTAAACAGGGGTGCCATATCAAG-----ATCT 510
QY 481 TTGTTTGGAGAGAAACTATTCAGCATCTAGAACAGAAATATGTTGCTCAACG 540
DB 511 TCATTTGAGCAACACCGCACTTCACGAGTGTGAGCAAAATTTGCTCCAGTGGTTCT 570
QY 541 AGAATAAATCTTCCCGTGCCTCCCAAGGACCCAGTCCAGGTTTATAAGCCACAGCAAAA 600
DB 571 AGGAGCCCAACCGAGCACCACATCAGCCAAAGACACACATCATGTTCCCAAGTTTCA 630
QY 601 ATGGAGTCTGCAAGTACTCTCAGTAGAGGGGAGAAACAGCA---GTAATTTGTCACCA 657

DB 631 CTCTCTCAACCCGAGAAATACCTCTTAATGAAACACACCACCAAAAGGTTTGAGAGGCCATCA 690
QY 658 GCATCAGATCTCTCAAAAGTGAGATTTTCTACTGATCTGTTTAAACATGCTATCAATGGAT 717
DB 691 GTTGCAACCACTCTTAAGTTGACTATGCACTGATCTCTTCAACATGTTATCAATGGAT 750
QY 718 GATTGCACTACAATACTCTCAGAGGCAACTCTCTGCGGATCTCTCTGCGGATGATCACTCA 777
DB 751 GGAACAAACCGAGAAAGAGCAGAGTCA-----TCTTCTAATGATGATAGTCA 798
QY 778 TGGCTGGCTTTTCTAGTCTGCTGGAGTGGTCAAAAGCGCAGAGAAATTTGTACAGCCCAAG 837
DB 799 TGGAGGGCTTTCCAGTCTGCGAGAGCCAGTACTAGTCTCAGACAAAAGGATTTCTCTAAG 858
QY 838 CTCTGCTGAGAGCAGTCTCTCTCCAGCTTCATCTTCTGAC-----TTTGAGGATTTGTT 891
DB 859 CCAGTAGAAGTAGCCCACTCTACATCAGGCATAGAGGACTTCTTTAAGATTCACCA 918
QY 892 AAGGACACACCTAATTTAACTCAACAGCACAAGATGTGAAGGGGATATCATG 951
DB 919 GCTGTGACAGTATCTCTAGCTCCAGCTGCCCCACAAAGTAAATGTGAAGATGATATCATG 978
QY 952 AGCTGTTTGGAGAGAGCAATATAGTATCGCTTTTGGCCATGTCATCAGCAACAGCTTGTCT 1011
DB 979 AGTCTGTTTGAAGTCCAGTATGTTTGGCCATATGCTGTCCAGCAGCAGCAACTTGTCT 1038
QY 1012 ATGCTCGCTCAGCAGCA 1028
DB 1039 TTTATGACCCCAAGCA 1055

AK068892 2157 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone: J023001G22, full insert sequence.

RESULT 8
AK068892
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team.; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group., Otsu, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsumoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kuwumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN.; Hara, A., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Itoh, M., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

1
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 2157)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiroka, T., Hori, F., Hotta, I., Iida, J., Ikeda, R., Iida, Y., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsuura, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nishikawa, J., Nishi, K., Oka, M., Ooka, H., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, B., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikawa, J., Oka, M., Ryū, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsuura, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hori, P., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

COMMENT

FEATURES
source

1..2157
/organism="oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023001G22"

ORIGIN

Query Match 21.1%; Score 307; DB 8; Length 2157;
Best Local Similarity 59.2%; Pred. No. 4.2e-74;
Matches 614; Conservative 0; Mismatches 390; Indels 33; Gaps 4;

QY 1 ATGACGAGAGAGCCCTCTCTAGGAGCTTAATGCCCGCATAGAAAGATTCTTGA 60
|||||
DB 277 ATGACGAGAGAGCCCTCTCTAGGAGCTTAATGCCCGCATAGAAAGATTCTG 336
|||||

QY 61 GGGCTCTTTAAACATCCAGAGACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCCAAGA 120
|||||
DB 337 GGCCTCTCGGCTGCCTGAGATAGAGATGTGCAGACTGCAAGTCAAGGGTCTCTGA 396
|||||

QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
|||||

DB 397 TGGCAAGTGTAAATCTTTGGATCTTTATATGCAATGCAAGTGTCTGGAAATCATAGAAC 456
QY 181 CTCGGGGTACACATATCAAGAGGTTTCGATCTGCCACTCTGGACACATGCTCCCCGAGCAG 240
|||||
DB 457 CTCGGGGTACATATATCTTAAGTAAAGTCCGCTACCTTGGATACATGCTGCCAGGACAA 516
|||||

QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
DB 517 GTTGCAATTTATCAATCAATGGGAAATGAAAAGTCAAAATAGTACTTGGGAAGCAGAGCTG 576
|||||

QY 301 CCCCCAACTATGATAGATGGAATTTGAGATTTTATACGTGCAAAAGTATGAGAGAG 360
DB 577 CCCCCTAACTATGATAGATGGAATTTGAGATTTTATACGTGCAAAAGTATGAGAGAG 636
|||||

QY 361 AGATGGGGTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGAGTCGAGCAGCAAGCGCGG 420
DB 637 AGATGGATACCGAGGAATGGGTCAATCAAGACCATCATCTGGTCTAGAGATGAGAGAGC 696
|||||

QY 421 AAATCTGTGGAGAGAGTGGGCCCGGGATATAGAGCATATGAGTAGTAGTCTGCTGTAAT 480
DB 697 TCAGAGTCACAAACTAGTGTGTTAAACAGGGGTGGCCATAATCAAAG-----ATCT 744
|||||

QY 481 TTGTTTGGAGAGAGAAAACCTATTTCAGCATCTAGAACAGAAATATGTTGCTGCAAG 540
DB 745 TCATTTGAGCAACACCGCCTTCCAGGAGTCAAGTGGAGCAAAATTTGCTCCAGTGGTTCT 804
|||||

QY 541 AGAATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCAAGTTATAAAGCCACAGCAGAAA 600
DB 805 AGGACGCCACCCAGGACCCACATAGCCAAAGCACAACCATCAGTTCCTCCAGAGTTTCA 864
|||||

QY 601 ATGAGTCTGCAGCTATCTCCAGTAGAGAGGAGGAGAAACAGCA---GTAAATGTTGCACA 657
DB 865 CGTCTCAACCCGAGAAATCACCTCTTAATGCAACACCCAAAGGTTGAGAGGCCATCA 924
|||||

QY 658 GCATCAGATCTCCAAAGTGATTTTGTACTGATCTGTTTAAACATGCTATCAATGAT 717
DB 925 GTTGACCACTCTTAAGTTGATCTGCAACTGATCTCTTCAACATGTTATCAATGAT 984
|||||

QY 718 GATTTCGACTACAAATACCTCAGAGGCAACTCTCTGGCGATACCTCTGGCGATGATAACTCA 777
DB 985 GGAACAACCGAGAAAGAGCAGAGTCA-----TCTTCTAATGATGATAGTAGTCA 1032
|||||

QY 778 TGGGCTGGCTTTTCACTGCTGGAAGTGGTCAAAAGGAGAGAAATTTGTACAGCCCAAG 837
DB 1033 TGGGAGGGCTTCCAGTCTGCAGAGCCAGTACCTAGCTCAGACAAAAGGATTTCTGCTAAG 1092
|||||

QY 838 CCTCTCAGAGCAGTCTCTCTCCAGCTTCTATCTTCTGAC-----TTTGAGGATTTGTT 891
DB 1093 CCAATAGAAAGTAAGCCCGAGTCTACATCAGGATAGAGGACTTGTTTAAAGATTCACCA 1152
|||||

QY 892 AAGGACACACCTAATTTAAACAACCTCAACAAGCACCACCAAGATGTGAAGGCGCATATCATG 951
DB 1153 GCTGTGACAGTATCTCTCAGCTCCAGCTGCCCGCCACAAGTAAATGTGAAGATGATATCATG 1212
|||||

QY 952 AGCTGTTTGAAGAGAGCAATATAGTATCGCTTTTGGCATGCAATCAGCAACAGGTTGCT 1011
DB 1213 AGTCTGTTGAGAAAGTCCAGTATGTTTTCGCCATATGCTGCTCCAGCAGCAGCAACTTCT 1272
|||||

QY 1012 ATGCTCGCTCAGCAGCA 1028
DB 1273 TTTATGACCCCGCCAGCA 1289
|||||

RESULT 9
LOCUS AB010695/1
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone MDK4.
ACCESSION AB010695
VERSION AB010695.1
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	CDS	GEKVRVFLKLGILDVHAPT" join(9646..9723,9907..10026,10127..10259,10309..10440, 10517..10939) /note="gene_id:MDK4.2 unknown protein" /pseudo /codon_start=1 /evidence=not experimental 11374..12294 /note="unnamed protein product; gb AAF34842.1 gene_id:MDK4.3 similar to unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAB10744.1" /db_xref="GI:9759494" /translation="MSVQNQNQLSVEPKSKRIDSEINSSSSNTCDHFFVYGCICN CRNVRHRSFDYLDVGLQSLDIATVTKRVTQTICFNDKKLHLVLDLHLLHT VMISNLTKEYLIEEDSRDLRLNGYSSSEFLIKLRPFVHEFLKANKRMFMYVI TMGRDYAMVNLNIDPEKYFGDVRITRNESFYIKTLDLADCEGCVIYDDTPHVM PDHKNLEITKNYFSDKTRHDVKYTKSYAEKRDSESRNDGSLANLVKIKQVYEGF FSGGVKDLIDSKDYLRLLLHDARRPKFYC" join(14111..14183,14300..14517) /note="unnamed protein product; gene_id:MDK4.4 unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAB10745.1" /db_xref="GI:9759495" /translation="MGSSKLLVAFLLIMVTISYDLTGIGIDARTVPTCYESCNAT FQNECNKMCVGLAYKDGSCYPPPEVDGLPKRPYFPCCNPIILSPSP" complement(join(19119..19815,20148..20277,20475..20607)) /note="gene_id:MDK4.5" /codon_start=1 /evidence=not experimental /product="Myb-related transcription factor-like protein" /protein_id="BAB10746.1" /db_xref="GI:9759496" /translation="MGKSSSESESEVKKQPTPEDEKLVGIOTHGPKWRTLPKNA GLKRCGSKRLRWNTYLRPKRGESLQSEETIQLHRLGNKWSAIAIHLPGRTDN EIKYNWNTHKKLLRMGIDPVTHCPRLNLLQSLFSLTSSLFKMSQPMWTFDLTTS NINPDILNHLTASLNNTYQNPQOQLNDLTQDTFTGLLNTSTPPVQWNGEYL GDYHSYTGDPSPNNKVPQAGNYSAAAFVSDHINDGNFKAGWFSMLAGTSSSS TPLNSSTFVYNGGSEDDRESFGSDMLMFFHHDDNNALNS" complement(join(26101..26460,26735..27223)) /note="unnamed protein product; gene_id:MDK4.6 pir t05759 strong similarity to unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAB10747.1" /db_xref="GI:9759497" /translation="MAPRLTCFVGKSGSTSTTTGKIGISGHNAANVADVPVGDGPIL VQLSSQCKTSPEAEMLVSRGLGDFDKIRRDGSSGSPANVLVHFVDYVWYDLGWLK DPYSGQCKTQKAYEIALNQDMFTFPVQVQVOLLGNEEELLKSVIAPRPSP AFRTAFORTSEYKATQALRMKVDSSGIDINVALYENGLENCARGENSGRVLNS DYVRKLEKLCVTKDLTGKXTVSTAHTFTWDFGNSKCGVAFVQLNTSLQIFGTQSP QLPDEI" complement(join(28446..28599,28832..29196,29427..29513, 29616..29852,29936..30044,30279..30592,31260..31475, 32131..32733)) /note="gene_id:MDK4.7" /codon_start=1 /evidence=not experimental /product="cyclic nucleotide and calmodulin-regulated ion channel" /protein_id="BAB10748.1" /db_xref="GI:9759498" /translation="MATEQEFTRASRRSSSVGYSEEDNTEDEEEMERIEE EEEEEDPRIGLTCGGRNGNSKMMGLRILDRSKWVRNKKVFLLVCATGL EVDPLFYLSVSDTCMLLDGMLATVTRALSMRTDLHLNWTIOPKIRRPYPG MONPDHNAAIRVVRVLYKSNPGGILLSTLTQINGWDSYAAACPLTRSLTQF FVQLGDTPLSKTKQPTVSRSSAEYRAMAFLTQELMWLKLVDLGVSHVQAMRI LVSVDLTVTFQYLPFKTVISRHURNASGXIIFGTVMGIALNMIAFYAAHAHA
TITLE	Kaneko,T., Kotani,H., Nakamura,Y., Sato,S., Asamizu,E., Miyajima,N. and Tabata,S. Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence features of the regions of 1,381,565 bp covered by twenty one physically assigned P1 and TAC clones	CDS	
JOURNAL MEDLINE PUBMED	DNA Res. 5 (2), 131-145 (1998) 9679202	CDS	
REFERENCE AUTHORS	2 (bases 1 to 78596) Nakamura,Y. Direct Submission	CDS	
TITLE	Submitted (28-JAN-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)	CDS	
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MDK4 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3//), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlini.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K18G13 and the 3' clone is GA469. Location/Qualifiers 1..78596 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="5" /clone="MDK4" /clone_lib="Mitsui P1" /ecotype="Columbia" complement(join(4967..6553,6955..8195,8901..9402)) /note="gene_id:MDK4.1" /codon_start=1 /evidence=not experimental /product="retroelement pol polyprotein-like" /protein_id="BAB10743.1" /db_xref="GI:9759493" /translation="MADAPPPTSPVIEVRTISPYDLTAADNSGAVISHPIKTNVNE EWACGFTALRSKRGFELDTGTPLOLPGSDPLEDLWLTINALLVSMWKTIDSELLTN ISHRDVADLMEQIRRFVSNQKQKADLTCQEGMTGMEYGLKLNKINDIN SYRPLRICASHMTGNELLSDMRSMSPVLIADGNKRVAVSEGTVRGLSHLILSV FYVKEGLDLSVGQMDENHCVOLADHFLVIOQRTTRWVTGIGKENGSCFPRGME NAAA VHTSKAPFDLWRLRHGASDKIVNLLPRELLSGKEILENVCDTCMRKQTRD TFPLSDNMSQSPQLHCDVMPGYPVTPSYSGARYFLTIVDDYSGVWVYVLTMDKSTQ RHLKDFIALVERQFTEIKTVRSNDGTFLCMRYEFLHGIITHEISCVGTPHQGRVE RKRHRLNLRALRFQSYLPQIFWGECILSAAVILNTPSMLQKGSFYEMLYKTPN YSHLRFVGLSCYAHNQNHKQKFAVRSRCVFPVPHGQKWLFLDEEQFPFVSRDV IFOETFPYKSMSCMRFTSSHKAFIAVTAAGTEPTTYNEAMVDKAWREAMEAIESLR VNOTESIYNLPGKRALGNKVYKIKYRSDGAIBRYKARLVVLGNCKQEGYDYDETFE PVAKSTVRLFLGVAAARDWVHMDVNAFLHGLDLEEVYMKLPQGFQCDPQKVCVR LHKSILYGLQPCWFSKLSALKQYGTQSLSDYSLSFYNNDGVFVHLVYVDDLLI SGSCDPAVAFQFSLSCFHNKDLGLLKYFLGIEVSRNAQGYLSQRYVLDIISMG LIGARPSAFPQKHLISLSPSSLSRYYRLVRLIYLAIVTRPESLYTAQF MONPDHNAAIRVVRVLYKSNPGGILLSTLTQINGWDSYAAACPLTRSLTQF FVQLGDTPLSKTKQPTVSRSSAEYRAMAFLTQELMWLKLVDLGVSHVQAMRI FSDSKSALISVNPVOHERTKHEVDCHFDIRDAILDGLIATSFPVPSHQLADILTKAL		
FEATURES source		CDS	

CDS
ACWYLLGVQSRACKLEOCENTIGCDRLMLSCBKYVYGTGTVMLDLRLAWAQNHOA
RSVCLDINTNYGTAYQWNTQIQLVSSSESLKILFPIFWGLMTLSTFGNLESTWSEV
VFNIIVLTGSLGLVMTWNTQIQLVSSSESLKILFPIFWGLMTLSTFGNLESTWSEV
VNYERORWAAMEGVDECEWVONLEGLARDIKYHLCLDLVROVPLFOHMDLVLNI
CDRUKSLITFKGETTQKEDGAQVRLMFPVVRGHLQSOLLRDGVKSCMLPGNPFSGDE
LUSWCLRRFPDPPSSSLVLTLETFAGGLDAEDDKYVTOHFRITVFNKVRKASR
YSPGWRWAAVAQVLAWRMYKRLRLTSLSLRPRRPLSRCSAGEBKRLRYAAILT
SPKPNDDPDDY
join(35464..35471,35682..35817,36015..36172,36315..36405,
36528..36647,36728..36796,36865..36956,37045..37147,
37247..37354,37476..37538,37610..37681,37785..37862,
38060..38122,38211..38271,38753..38835,38979..39092,
39212..39289,39376..39472,39534..39766,39879..40002,
40161..40364,40479..40511)
/note="emb|CAB50793.1
gene_id:MDK4.8"
/codon_start=1
/evidence=not_experimental
/product="DNA repair and meiosis protein Mre11"
/protein_id="BAB10749.1"
/db_xref="GI:9759499"
/translation="MSREDFSDTLRLVATDCHLGMKDEIRRHDSFKAPEEISIA
EKQVDFLLGLDGLHENKPSRTLVKAIKILRRHLNDKPKQVQVSDQTVNFOAF
GQVNEDEPHNVGLPWFVSHGNHDDPAGVDNLISADILSACNLVNFKQNLGSSGVG
QITLYPILMKGSTVALYGLGNIRDERLNRFQPHAVQWRPVPVQSGCDVDFWFI
LVLHQRVSNPKNAISEHFLPFLDFIVMGHEHECLIDPQEVSMGSHITQPGSSVA
TSLIDGESKPHVLLLEIKGNOYRPTKPLTSVRPFVETIIVLDESIDPNDQNSIL
EHLDKVRLNLEKASKANRSEIKLPLVIRKVDYGFMTINPQKQYKVKVNPQ
DILIPSKSKGRSEANIDSELRPEELNQONIRALVAESNLKWEILPVNDLVALH
NFMKDDKLAFSCVQVNLQETRGKLAKSDAKKEEDDLILKVGECLEERLKORSTR
PTGSSQFLSTGSENLTGSGSIANASFDDEDITQMSGLAPPRGRSGSTEDDVS
GKAPTRGRGRGKSSAMKQTLDSLSGFRQSRQSSASAAAFKSASTIGEDDVS
PSEVEPDDFNKPPSSDDESTGKGRKRPATTKRGRSGSTSKRGRKNESSSL
NRLSSKDDDEDDREKLNKOPVRVNYGALRR"
join(41186..41311,41410..41571,41636..42145)
/note="gb|AD28773.1
gene_id:MDK4.9"
/codon_start=1
/evidence=not_experimental
/product="LhcB3 chlorophyll a/b binding protein"
/protein_id="BAB10750.1"
/db_xref="GI:9759500"
/translation="WASTPTSSSVLTPTFLFGQTKASSFNPLRDVSLGSPKPYTMGN
DLWYGPDRVYKLGVSQVTPSYLTGFGDYGMDTAGLSADPEAFKRALEVIHGRW
AMLGAFGCTTPVLQKVRVDFKPEVPMFKAGSQIFSEGLDYLGNPLVHAQSILAVL
GFQVILMLGLVEFRINGLDVGEENDLYPGGYFDPGLGLADDPVTFABLKVKKEIKNGR
Query Match 17.6%; Score 255.4; DB 8; Length 78596;
Best Local Similarity 60.1%; Pred. No. 1.2e-59;
Matches 877; Conservative 0; Mismatches 1; Indels 581; Gaps 4;
QY 575 GTCAGGTTATAAGCCACAGCAGAAAATGGAGTCTGCAGCTACTCCAGTAGAGGGAGA 634
DB 61482 GACAGGTTATAAGCCACAGCAGAAAATGGAGTCTGCAGCTACTCCAGTAGAGGGAGA 61423
QY 635 AACAGCAGTAATGTTGCACAGCATCAGATCCTCCAAAGTGGATTGTTGCTACTGATC 694
DB 61422 AACAGCAGTAATGTTGCACAGCATCAGATCCTCCAAAGTGGATTGTTGCTACTGATC 61363
QY 695 TGTTTAACTGCTATCAATGGATGATTGCTACTACAAATACCTCAGAGGCAACTCTGGCG 754
DB 61362 TGTTTAACTGCTATCAATGGATGATTGCTACTACAAATACCTCAGAGGCAACTCTGGCG 61303
QY 755 ATACTCTCCGATGATACTCATGGCTGGCTTTTCAGT----- 793
DB 61302 ATACTCTCCGATGATACTCATGGCTGGCTTTTCAGT----- 793
QY 794 ----- 793
DB 61242 GATACTAACTCGGGCAGACATTTTTCCTATCTCTGTTTAACTTTATGTCATTTGTTCT 61183
QY 794 -----CTGCTGG 800
|||||

DB 61182 CTTTGTGAAAATGCTTTTGAGGTAAATAGATTAATTTCTGAATCTCTAAACCTGTAGCTGCTGG 61123
QY 801 AAGTGGTCAAAACGGCAGAGAAAATTTGTCAAGCCAAAGCTGTGAGAGCAGTTCTCTCTCC 860
DB 61122 AAGTGGTCAAAACGGCAGAGAAAATTTGTCAAGCCAAAGCTGTGAGAGCAGTTCTCTCTCC 61063
QY 861 AGCTTCATCTCTGACTTTGAGGATTTGTTTAAAGGACACACCTTAATTTAACTCAACA 920
DB 61062 AGCTTCATCTCTGACTTTGAGGATTTGTTTAAAGGACACACCTTAATTTAACTCAACA 61003
QY 921 AGCACCACAAAGATGTGAAGGCGGATATCATGAGCTGTTTGCAG----- 964
DB 61002 AGCACCACAAAGATGTGAAGGCGGATATCATGAGCTGTTTGAAGAGGTAATATTAGTT 60943
QY 965 ----- 964
DB 60942 AATTATATACGCTTCTCAGATTTTCAGTCTTGTGTTGATATTTCAGCTTTTCCAAGTTTCGTT 60883
QY 965 ----- 964
DB 60882 TTTGTATCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 60823
QY 965 ----- 964
DB 60822 TATCACTTATCATGATTTCTTCAACACGGGACATTTGGACACACGAGCAGATTCAAGTG 60763
QY 965 -----AGACGAATATAGTATGCTGCTTTTGCCTTTTGCATGCA 995
DB 60762 ATATAACTCTGTGTGTTCTTATTTGATTACAGACGAATATAGTATGCTGCTTTTGCATGCA 60703
QY 996 TCAGCAACAGTTGCTATGCTGCTGCTCAGCAGCAAGCCCTTTTACATGGCTGCGAGCAAGC 1055
DB 60702 TCAGCAACAGTTGCTATGCTGCTGCTCAGCAGCAAGCCCTTTTACATGGCTGCGAGCAAGC 60643
QY 1056 TGCTGGAGGCACTCAAAACGGCTGAATCAACAAAGCTATTGCTAATGCTCTTAACTGATG 1115
DB 60642 TGCTGGAGGCACTCAAAACGGCTGAATCAACAAAGCTATTGCTAATGCTCTTAACTGATG 60583
QY 1116 TTCTGCAAAATTTGGTCAAAACGGCTGCTACAGATCCCGGAATGATTAACCCCGTAGG 1175
DB 60582 TTCTGCAAAATTTGGTCAAAACGGCTGCTACAGATCCCGGAATGATTAACCCCGTAGG 60523
QY 1176 TGCTCAAGCTGATCTCCAGAACTTATGCA----- 1206
DB 60522 TGGTCAAGCTGATCTCCAGAACTTATGCA----- 1206
QY 1207 ----- 1206
DB 60462 AAGATCAATATTACCATTTGACCAAGTTCTTTCGAAAATTTCTCACTAATTTCTTGTCTCTG 60403
QY 1207 -----AACATGAATATGAACGCAAAACATGAACACGAGACCCGAC 1246
DB 60402 TTACATCGCCATAATTTTAGAACATGAATATGAACGCAAAACATGAACACGAGACCCGAC 60343
QY 1247 AACCGCAGAGAAACACTCTCAATACCCATCATC----- 1280
DB 60342 AACCGCAGAGAAACACTCTCAATACCCATCATC----- 1280
QY 1281 ----- 1280
DB 60282 GTTTGTTTCTTCTGAAAACATCAACGTTATCAAAAGTGTCTTAACTTAACTTAACTTCA 60223
QY 1281 -----CAGTTTCTCAACAATGGGTCAAGCTAATCAAGT 1313
DB 60222 TATGATCTTTGGTGTGTTTAAATGAGTCTTCTACAAATGGGTCAAGCTAATCAAGT 60163
QY 1314 GAAAGGATGACCCCAAACTCAACGGGTAAACCTCAGTCATCATCCGCAACCCCAACCAAC 1373
DB 60162 GAAAGGATGACCCCAAACTCAACGGGTAAACCTCAGTCATCATCCGCAACCCCAACCAAC 60103
QY 1374 AAGCACCACACATCTTCCAAATCAGGCAAAAGATTTTGAATTTCTTCTCTTGTGATGGTGG 1433
DB 60102 AAGCACCACACATCTTCCAAATCAGGCAAAAGATTTTGAATTTCTTCTCTTGTGATGGTGG 60043

```

QY 1434 AATGTTTCAAAACATTGA 1452
Db 60042 AATGTTTCAAAACATTGA 60024

RESULT 10
LOCUS CO455437 371 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 21197 from Patent WO0192523.
ACCESSION CO455437
VERSION CO455437.1 GI:41421949
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets, R.A. and Leach, M.D.
TITLE Human polynucleotides and polypeptides encoded thereby
JOURNAL Patent: WO 0192523-A 21197 06-DEC-2001;
Curagen Corporation (US)
FEATURES
source
1..371
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 16.3%; Score 237.2; DB 6; Length 371;
Best Local Similarity 81.4%; Pred. No. 1.1e-54;
Matches 275; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGAGCTTAATGCCCGCCATAGAAAGATTCTTGA 60
Db 34 ATGAACGAGAGGCGTCCGCTCTCAAGGAGCTCAACGCCAAGCACACAGAAATATTGAA 93
QY 61 GGGCTCTTAAACATCCAGACAGACAGAGATGCTGACTGCAAAACAAAGAGTCCCAAGA 120
Db 94 GGTCTTCAACGGCATCTGAGAAATAGAGAAATCGCAGACTGCAAGTCAAGGGTCTCTGA 153
QY 121 TGGGCTAGTGTAAATTAGGTATCTTATCTCATGCAATGTTCTGGGATTCACAGAGT 180
Db 154 TGGCAAGTGTGAATCTAGTATCTTATATCATGACACATGTTCTGGCATTATAGAGC 213
QY 181 CTCGGGGTACACATATCGAAGTTGATCTGCACTCTGGAACATGCTCCCGAGCAG 240
Db 214 CTGGGGTGCATATCTAAGTAAAGTCTGCCACCTGGATACATGCTGCCAGCAA 273
QY 241 GTTGCAATTATACAGTCAATGGGAATGATTAAGCAATAGTTACTGGGACGAGCTA 300
Db 274 GTTGCAATTATTAATCAATCGGAACGAAAGCAAAATAGCTATTTGGGAACGAGCTG 333
QY 301 CCCCAACTAGATAGAGTTCGAATTGGAATTTAT 338
Db 334 CTCCTAACTACGATAGGGTTGGAATAGGAATTTGAT 371

RESULT 11
LOCUS AR245215 270 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 574 from patent US 6476212.
ACCESSION AR245215
VERSION AR245215.1 GI:27293089
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1
AUTHORS Lalgudi, R.V., Ito, L.Y. and Sherman, B.K.
TITLE Polynucleotides and polypeptides derived from corn ear
JOURNAL Patent: US 6476212-A 574 05-NOV-2002;
FEATURES
Location/Qualifiers

```

```

source 1..270
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 13.7%; Score 199.6; DB 6; Length 270;
Best Local Similarity 83.7%; Pred. No. 3.4e-44;
Matches 226; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 92 GTGCTGACTGCAAAACAAAGGTCCAAGATGGGCTAGTGTAAATTTAGGTATCTTTATCT 151
Db 1 GCGCAGACTGCAAGTCAAGGTCCTCGATGGGCAAGTGTGAATCTAGGTATCTTTATAT 60
QY 152 GCATGCAATGTTCTGGGATTCACAGGAGTCTCGGGTACACATATCGAAGTTCGATCTG 211
Db 61 GCATGCAATGTTCTGGCATTTCATAGAAGCTCGGGGTGCACATATCTAAGGTAAGATCTG 120
QY 212 CCATCTGGACACATGCTCCCGAGCAGGTTGCAATTATACACATCAATGGAATGATA 271
Db 121 CCACCTGGATACATGCTGCCAGAGCAAGTTGCAATTTATTTCAATCAATGGAACGAAA 180
QY 272 AAGCAATAGTTACTGGGAGACAGAGCTACCCCAAACTATGATAGAGTTCGAAATTGAGA 331
Db 181 AAGCAATAGCTATTGGAGACAGAGTCCCTCTACTAGTAGGTTGGATGATAGAGA 240.
332 ATTTTATACATGTCGAAAGTATGAGAGAGA 361
241 ATTTATCCCTGGCAAAATATGAGGACAAGA 270

RESULT 12
LOCUS BC077937 3504 bp mRNA linear VRT 20-SEP-2004
DEFINITION Xenopus laevis MGC80897 protein, mRNA (cDNA clone MGC:80897
IMAGE:5515434), complete cds.
ACCESSION BC077937
VERSION BC077937.1 GI:50417733
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 3504)
AUTHORS Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
Dev. Dyn. 225 (4), 384-391 (2002)
JOURNAL 12454917
PUBMED 2 (bases 1 to 3504)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klauser, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Brange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL 12477932
PUBMED

```

REFERENCE	3 (bases 1 to 3504)	
AUTHORS	Klein, S. and Gerhard, D.S.	
TITLE	Direct Submission	
JOURNAL	Submitted (19-JUL-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA	
REMARK	NIH-MGC Project	
COMMENT	Contact: XGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Dr. Igor Dawid cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca	
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov</p> <p>Series: IRAC Plate: 155 Row: 9 Column: 13</p> <p>This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.</p> <p>Location/Qualifiers</p> <p>1..3504</p> <p>/organism="Xenopus laevis"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:8355"</p> <p>/clone="MGC:80897 IMAGE:5515434"</p> <p>/tissue_type="Embryo, stage 31/32, Xenopus"</p> <p>/clone_lib="NICHD_XGC_Emb4"</p> <p>/lab_host="DH10B"</p> <p>/note="Vector: pCMV-SPORT6"</p> <p>1..3504</p> <p>/gene="MGC80897"</p> <p>/db_xref="LocusID:446881"</p> <p>384..1436</p> <p>/gene="MGC80897"</p> <p>/codon_start=1</p> <p>/product="MGC80897 protein"</p> <p>/protein_id="AAH77937.1"</p> <p>/db_xref="GI:50417734"</p> <p>/db_xref="LocusID:446881"</p> <p>/translation="MTTSRERKAQKLNHQHQAIIISRLREEDNKYCADCEAKGPRWA SNLGVFCIRGAGIHRNLGVHISRVKSVNLDQWTPQIQCMQDMGNTKARRIYEANL PENFRPQDSVEFFIRDKYRKRYDKNATVSGTKEKQKKEERKARPEKRWKP VGADVLKKEIPLVKASPKPSKPTVLLGLDAPVEAPALNGTSKTAARVNDLDFI GMVSNPIPVSAAPQVTSNPASTSVSSSSDLDTDDTTKEDTQKPLKSDSI LSLYTANQQQNTTGMFMGSSPFPYSQPSASFCAPFMSGVPIAAGGIMGNMGPSV PMGTGECDDGNGRIA"</p>	
gene		
CDS		
ORIGIN	<p>Query Match 7.7%; Score 111.2; DB 5; Length 3504; Best Local Similarity 59.5%; Pred. No. 1.7e-19; Matches 188; Conservative 0; Mismatches 128; Indels 0; Gaps 0;</p>	
QY	23	CTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAAGGGTCTTTAAACATCCAGAGA 82
Db	412	CTCAGAAGCTGAATGAGCAGCATCAGGCATCTTGTCCCGATGCTTCGGGAGGAAGACA 471
QY	83	ACAGAGATGCTGACTGCTCAAAACAAAGGTCGAGATGGGCTAGTGTAAATTAGGTA 142
Db	472	ACAAGTATTGTGCTGACTGCTGAAGCAAAAGGTCCTAGGTGGGCTTCTCGAACCTAGGAG 531
QY	143	TCTTTATCTGCATGCAATGTTCTGGGATTTCACAGGAGTCTCGGGGTACACATATCGAAGG 202
Db	532	TTTTATGTGCTTCGTTGTTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591
QY	203	TTGATCTGCCACTCTGACACATGCTCCCGAGCAGGTTCGTCATTATACAGTCAATGG 262
Db	592	TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGCAGGATATGG 651
QY	263	GAAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322
Db	652	GAAATACAAAGGCAAGAGCTATATACGAGGTATCTGCCAGAAAACCTTCGAGAGACCTC 711
QY	323	GAAATGAGAAATTTTAT 338
Db	712	AGACTGATCAATCTGT 727
RESULT 13		
BC073437		
LOCUS		
DEFINITION	Xenopus laevis hypothetical protein LOC43647, mRNA (cDNA clone IMAGE:5516053), partial cds.	
ACCESSION	BC073437	
VERSION	BC073437.1	GI:49116706
KEYWORDS		
SOURCE	Xenopus laevis (African clawed frog)	
ORGANISM	Xenopus laevis	
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus; Xenopus.	
REFERENCE	1 (bases 1 to 778)	
AUTHORS	Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W. and Richardson, P.	
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative	
JOURNAL	Dev. Dyn. 225 (4), 384-391 (2002)	
PUBMED	12454917	
REFERENCE	2 (bases 1 to 778)	
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, K.J., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PUBMED	12477932	
REFERENCE	3 (bases 1 to 778)	
AUTHORS	Klein, S. and Gerhard, D.S.	
TITLE	Direct Submission	
JOURNAL	Submitted (14-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA	
REMARK	NIH-MGC Project	
COMMENT	Contact: XGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Dr. Igor Dawid cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre,	

BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca		BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca	
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Maasson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliqne Schnerch, Ursula Skaleka, Duane Smalish, Jeff Stott, Miranda Tsai, George Yang, Jacqueline Schein, Asim Siddiqui, Rob Holt, Marco Marra.		Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Maasson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliqne Schnerch, Ursula Skaleka, Duane Smalish, Jeff Stott, Miranda Tsai, George Yang, Jacqueline Schein, Asim Siddiqui, Rob Holt, Marco Marra.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 155 Row: k Column: 7 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 155 Row: k Column: 7 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.	
FEATURES	source	FEATURES	source
	1..778		1..554
	/organism="Xenopus laevis"		/organism="Homo sapiens"
	/mol_type="mRNA"		/mol_type="mRNA"
	/db_xref="taxon:8355"		/db_xref="taxon:9606"
	/clone="IMAGE:5516053"		/clone="IMAGE:3856134"
	/tissue_type="Embryo, stage 31/32, Xenopus"		/tissue_type="Ovary, adenocarcinoma"
	/clone_lib="NICHK_XGC_Emb4"		/clone_lib="NIH_MGC_66"
	/lab_host="DH10B"		/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"		/note="Vector: pCMV-SPORT6"
gene	1..>778		1..>554
	/gene="LOC443647"		/gene="LOC443647"
CDS	394..>778		12477932
	/gene="LOC443647"		22388257
	/codon_start=1		22388257
	/product="LOC443647 protein"		22388257
	/protein_id="AAH73437.1"		22388257
	/db_xref="GI:49116707"		22388257
	/db_xref="LocustID:443647"		22388257
	/translation="MTTSREKAKLNEHQHILSRMLREDNDKYCADCAKGRPA SWNLGFCIRGAGHRLNIGVHSRKSVNLVDQWTEPQIQMDMGNTRQKRYANL QENFRPQDQSVFPIRDKYKKKK"		22388257
ORIGIN		REMARK	
Query Match	7.5%; Score 109.6; DB 5; Length 778;	COMMENT	
Best Local Similarity	59.2%; Pred. No. 4.5e-19;		
Matches 187; Conservative 0; Mismatches 129; Indels 0; Gaps 0;			
QY 23 CTAAGAGCTTAATGCCCCATAGAAAGATTCTTGAAGGGCTTCTTAACATCCAGAGA 82			
DB 422 CTCAGAGCTGAATGAGCAGCATCAGGCCATCTTGCCAGGATGCTCCGCGAGGATGACA 481			
QY 83 ACAGAGAAATGTCTGACTGCAAAACAAAGGTCCAGATGGGTAGTGTATTTAGTGA 142			
DB 482 ACAAGATTGTCTGACTGTGAGGCAAAAGGTCCTAGGTGGGCTTCTGGAACTTAGGGG 541			
QY 143 TCTTATCTGCATGCAATCTTCTGGGATTCAGAGGATCTCGGGGTACACATATCGAAGG 202			
DB 542 TTTTATCTGCATCTGTTGTGTGGAAATTCATAGAAACCTTGGAGTTTCATATCAAGG 601			
QY 203 TTCGATCTGCCACTCTGGACATATGCTCCCGAGCAGGTTCGATTTATACAGTCAATGG 262			
DB 602 TGAATCGGTAAATTTGGATCAATGACTCCCGACAAATTCAGTCATGCGAGGATATGG 661			
QY 263 GAAATGATAAGCAATAGTTACTGGGAAGCGAGGCTACCCCAACTATGATGATGATTG 322			
DB 662 GAAACAAGGCGACAGCAAGAAATAGAGCAAAATCTACAGGAAACCTTCAGAGACCTC 721			
QY 323 GAATTCAGAAATTTAT 338			
DB 722 AAATGATCAATCTGT 737			
RESULT 14			
BC008672			

LOCUS	BC008672	554 bp	mRNA	linear	PRI 16-SEP-2003
DEFINITION	Homo sapiens stromal membrane-associated protein 1, mRNA (CDNA IMAGE:3856134), partial cds.				
ACCESSION	BC008672				
VERSION	BC008672.2 GI:34189698				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (Bases 1 to 554)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Dietzenko, L., Narusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skaleka, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PubMed	12477932				
REFERENCE	2 (Bases 1 to 554)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	On Aug 25, 2003 this sequence version replaced gi:16306911. Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/BTP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseghe, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 13 Row: d Column: 9 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21264557.				
	Location/Qualifiers				
	1..554				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:3856134"				
	/tissue_type="Ovary, adenocarcinoma"				
	/clone_lib="NIH_MGC_66"				
	/lab_host="DH10B"				
	/note="Vector: pCMV-SPORT6"				
	1..>554				
gene	1..>554				


```

/genes="SMAP1"
/notes="synonyms: FLJ13159, SMAP-1"
/db_xref="LocusID:60682"
116...>554
/genes="SMAP1"
/codon_start=1
/product="SMAP1 protein"
/protein_id="AAH08672.1"
/db_xref="GI:34189699"
/db_xref="LocusID:60682"
/translation="WATRSCKRAQKLNQHOHLILSKLRREDNKYCADCEAKGPRWA
SWNGVFCIRCAIGHNLSRVKSVNLDQWTAQIQCMQDMGNTKARLLYEANL
PENFRPQTDQAVEFFIRDKYKKYDDKNAITNKEKKKKKKKKKKKKKKKK
LTAFLKQDQQLPEPKKSTSPKAAEPTVLLGLDPAVAVPTVNGTTPVPLNDLDI
FGPMISNPLPATVMPPAQGTSPAPAAATLSTVTSQDLPLPTEQTKSBEVAKQLSKD
SILSLYGTGTIOQSTPGVFMGPTNIPTSOAPAAFOGFSMGVPAAPGLIGNVMG
QSPSMVGMPPMPNGFMGNACTGUMPLQNVVVGPOGMVQMGAPQSKFGLPQAQPOW
SLSQMNQMGMSISSATPAGPGSPSTTAGWSGSSGQTLSTQLMK"
misc_feature
149..421
/genes="SMAP1"
/notes="COG5347; Region: GTPase-activating protein that
regulates ARFs (ADP-ribosylation factors), involved in
ARF-mediated vesicular transport [intracellular
trafficking and secretion]"
/db_xref="CDD:COG5347"

ORIGIN
Query Match 7.4%; Score 107.2; DB 9; Length 554;
Best Local Similarity 60.1%; Pred. No. 2.1e-18;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 23 CTAAGGAGCTTAATGCCGCCCATAGAAAGATTCTTGAAGGCTTCTTAAACATCCAGAGA 82
DB 144 CTCAGAGCTGAACGAGGAGCACCGCTCATCTTATCCAGCTTCTGAGGAGGAGGACA 203

QY 83 ACAGAGAATGTGCTGACTGCAAAACAAAGGTCCAGATGGGCTAGTGTAAATTAGGTA 142
DB 204 ACAAGTACTGCCGAGCTCGAGGCGCAAGGTCTCTGATGGGCTTCTTGAATATTGGTG 263

QY 143 TCTTTATCGATGCATAGTTCTTGGGATTCACAGAGTCTCGGGTACACATATCGAAG 202
DB 264 TGTTTATTGTCATCAGATGTCTGGAATTCATAGAAATCTTGGGGTTCATATATCCAGG 323

QY 203 TTCGATCTGCCACTCTGGACACATGGCTCCCGAGCAGGTTCGATTTATACAGTCAATGG 262
DB 324 TCAATCAGTCAACCTAGACCAATGGACAGCAGACAGATACAGTGCATGCAAGATATGG 383

QY 263 GAAATGATAAACAATAGTTACTGGGAGCAGAGCTACCCCAAACTATGATAGA 318
DB 384 GAAATACTAAAGCAAGACTACTCTATGAAGCAATCTTCCAGAGAACCTTTCGAAGA 439

RESULT 15
AY055004 2223 bp mRNA linear PRI 20-OCT-2001
LOCUS Homo sapiens stromal membrane-associated protein SMAP1B (SMAP1)
DEFINITION mRNA, complete cds, alternatively spliced.
ACCESSION AY055004
VERSION AY055004.1 GI:16303737
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2223)
AUTHORS Marcos,I., Borrego,S., Rodriguez de Cordoba,S., Galan,J.J. and
Antinolo,G.
TITLE Cloning, characterization and chromosome mapping of the human SMAP1
gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2223)
AUTHORS Marcos,I., Borrego,S., Rodriguez de Cordoba,S., Galan,J.J. and
Antinolo,G.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2001) Unidad de Genetica Medica y Diagnostico
Prenatal, Hospitales Universitarios Virgen del Rocío, Avda. Manuel
Siurot s/n, Sevilla 41013, Spain
FEATURES

```

```

source
1..2223
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
1..2223
/genes="SMAP1"
/genes="SMAP1"
/genes="SMAP1"
/genes="SMAP1"
/translation="WATRSCKRAQKLNQHOHLILSKLRREDNKYCADCEAKGPRWA
SWNGVFCIRCAIGHNLSRVKSVNLDQWTAQIQCMQDMGNTKARLLYEANL
PENFRPQTDQAVEFFIRDKYKKYDDKNAITNKEKKKKKKKKKKKKKKKK
LTAFLKQDQQLPEPKKSTSPKAAEPTVLLGLDPAVAVPTVNGTTPVPLNDLDI
FGPMISNPLPATVMPPAQGTSPAPAAATLSTVTSQDLPLPTEQTKSBEVAKQLSKD
SILSLYGTGTIOQSTPGVFMGPTNIPTSOAPAAFOGFSMGVPAAPGLIGNVMG
QSPSMVGMPPMPNGFMGNACTGUMPLQNVVVGPOGMVQMGAPQSKFGLPQAQPOW
SLSQMNQMGMSISSATPAGPGSPSTTAGWSGSSGQTLSTQLMK"
3'UTR
1392..2223
/polyA_signal
1677..1682
/genes="SMAP1"
/genes="SMAP1"

ORIGIN
Query Match 7.4%; Score 107.2; DB 9; Length 2223;
Best Local Similarity 60.1%; Pred. No. 2.2e-18;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 23 CTAAGGAGCTTAATGCCGCCCATAGAAAGATTCTTGAAGGCTTCTTAAACATCCAGAGA 82
DB 97 CTCAGAGCTGAACGAGCAGCACCGCTCATCTTATCCAGCTTCTGAGGAGGAGGACA 156

QY 83 ACAGAGAATGTGCTGACTGCAAAACAAAGGTCCAGATGGGCTAGTGTAAATTAGGTA 142
DB 157 ACAAGTACTGCCGAGCTCGAGGCGCAAGGTCTCTGATGGGCTTCTTGAATATTGGTG 216

QY 143 TCTTTATCGATGCATAGTTCTTGGGATTCACAGAGTCTCGGGTACACATATCGAAG 202
DB 217 TGTTTATTGTCATCAGATGTCTGGAATTCATAGAAATCTTGGGGTTCATATATCCAGG 276

QY 203 TTCGATCTGCCACTCTGGACACATGGCTCCCGAGCAGGTTCGATTTATACAGTCAATGG 262
DB 277 TCAATCAGTCAACCTAGACCAATGGACAGCAGACAGATACAGTGCATGCAAGATATGG 336

QY 263 GAAATGATAAACAATAGTTACTGGGAGCAGAGCTACCCCAAACTATGATAGA 318
DB 337 GAAATACTAAAGCAAGACTACTCTATGAAGCAATCTTCCAGAGAACCTTTCGAAGA 392

```

Search completed: September 27, 2005, 05:03:34
Job time : 6240 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 02:51:39 ; Search time 282 Seconds
(without alignments)
8425.084 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaaagccaaagct.....gaatgttcacaaacattga 1452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A-COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS-COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199.6	13.7	270	4	US-09-313-294A-574
2	182.4	12.6	408	4	US-09-640-211A-1383
3	107.2	7.4	2885	4	US-09-774-528-3
4	94.2	6.5	1447	3	US-09-484-970B-121
5	94.2	6.5	1976	4	US-09-620-312D-515
6	86.2	5.9	309	4	US-09-513-999C-2719
7	86.2	5.9	459	4	US-09-621-976-845
8	82.6	5.7	1327	4	US-09-620-312D-1104
9	78.2	5.4	1261	4	US-09-855-323-4
10	74.8	5.2	514	4	US-09-854-133-660
11	72.2	5.0	435	4	US-09-640-211A-1584
12	68.6	4.7	6990	4	US-09-620-312D-195
13	68.4	4.7	332	4	US-09-513-999C-384
14	65.4	4.5	2119	4	US-09-620-312D-35
15	63.4	4.4	7218	1	US-08-232-463-14
16	62.8	4.3	333	4	US-09-640-211A-1594
17	62	4.3	1455	4	US-09-248-796A-4726
18	55.6	3.8	286	4	US-09-313-294A-4791
19	55.6	3.8	543	4	US-09-248-796A-4610
20	49.4	3.4	1257	4	US-09-270-767-10482
21	47.8	3.3	505	4	US-09-513-999C-1131
22	47.8	3.3	1354	4	US-09-620-312D-804
23	47.4	3.3	213	4	US-09-513-999C-22310
24	46.6	3.2	502	4	US-09-828-303-3
25	46.6	3.2	2041	4	US-09-828-303-11
26	45.4	3.1	1141	4	US-09-806-708B-22
27	45.2	3.1	2712	4	US-09-023-905A-11

28	45.2	3.1	4595	4	US-09-023-905A-9	Sequence 9, Appli
29	44	3.0	5330	4	US-09-023-905A-1	Sequence 1, Appli
30	42.4	2.9	473	4	US-09-471-276-723	Sequence 723, App
31	42.4	2.9	562	4	US-09-621-976-338	Sequence 338, App
32	41.8	2.9	2949	4	US-09-023-905A-8	Sequence 8, Appli
33	41.8	2.9	5954	4	US-09-023-905A-6	Sequence 6, Appli
34	41.4	2.9	699	4	US-09-248-796A-5451	Sequence 5451, Ap
35	41.2	2.8	3456	4	US-09-023-905A-5	Sequence 5, Appli
36	41.2	2.8	4382	4	US-09-023-905A-3	Sequence 3, Appli
37	40.4	2.8	235064	4	US-09-949-016-15390	Sequence 15390, A
38	38.8	2.7	957	3	US-09-134-001C-1852	Sequence 1852, Ap
39	38.8	2.7	1194	4	US-08-710-279-1243	Sequence 1243, Ap
40	38.8	2.7	2998	4	US-09-710-279-3506	Sequence 3506, Ap
41	38.8	2.7	44342	4	US-09-949-016-12661	Sequence 12661, A
42	38.6	2.7	333	4	US-09-640-211A-1761	Sequence 1761, Ap
43	38.6	2.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
44	37.2	2.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
45	36.8	2.5	1830121	4	US-09-557-884-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-313-294A-574
; Sequence 574, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 574
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700549464HI
US-09-313-294A-574

Query Match		13.7%	Score 199.6;	DB 4;	Length 270;
Best Local Similarity		83.7%	Pred. No. 4.6e-57;		
Matches 226;		Conservative 0;	Mismatches 44;	Indels 0;	Gaps 0;
QY	92	GTGCTGACTGCAAAACAAAGGTCCAGATGGCTAGTGGTAAATTTAGGTATCTTTATCT	151		
Db	1	GGCGAGCTGCAAGTCAAGGGTCTCGATGGCAAGTGTGAATCTAGTATCTTTATAT	60		
QY	152	GCATCAATGTTCTGGGATTCACAGAGTCTCGGGGTGCACATATCGAAGTTCGATCTG	211		
Db	61	GCATGACATGTTCTGGCATTTCATAGAGCTCGGGGTGCACATATCTAGGTAAGATCTG	120		
QY	212	CCACTCTGCACATGCTCCCGAGCAGGTTCATTTATACAGTCAATGGGAAATGATA	271		
Db	121	CCACCTCGATACATGCTGCGCAGAGCAAGTTCATTTATTCATCAATGGGAAACGAA	180		
QY	272	AAGCAATAGTTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTGGAATGAGA	331		
Db	181	AAGCAATAGTTTGGGAGCAGAGCTCCCTCTTAACATGAGGTGGGATGAGATA	240		
QY	332	ATTTTATAGTGCAGAAAGTATGAAGAGAAGA	361		
Db	241	ATTTTCATCGTGCAAAATATGAGGACAAGA	270		

RESULT 2
US-09-640-211A-1383

```
; Sequence 1383, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
;   Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1383
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-09-640-211A-1383

Query Match      12.6%; Score 182.4; DB 4; Length 408;
Best Local Similarity 85.0%; Pred. No. 4.4e-51;
Matches 204; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 ATGACGAGAGAGCCAAAGCTCTTAAGGAGCTTAATGCCCGCCATAGAGAAAGATTCTTGAA 60
Db 169 ATGACGGGAGGCGCAAGCTCTCCAGGAGCTCAATGCCACCAAGAAAGATTCTCGNA 228

QY 61 GGGCTCTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAGGTCCAAGA 120
Db 229 GGGCTCTCAAAATGGCTGAGAACAGGGAGTGTGCTGATTGTAAGGCCAAAGGTCCAAGA 288

QY 121 TGGGCTAGTGTAAATTTAGGTATCTTATCTGCATGCATGTTCTGGGATTCACAGGAGT 180
Db 289 TGGGCTAGTGTCAATTTAGGATATTTATGATGATCAATGTTCAAGGAATCCATAGAAGT 348

QY 181 CTCGGGGTACACATATCGAAGGTTGATCTGCCACTCTGGACACATGCTCCCGGAGCAG 240
Db 349 CTTGGGGTACACATATCGAAGGTTCCGATCAGTCTACTTTGGACACATGCTTCCAGAGCAG 408

RESULT 3
US-09-774-528-3
; Sequence 3, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
;   Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 3
; LENGTH: 2885
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (120)...(1442)
US-09-774-528-3

Query Match      7.4%; Score 107.2; DB 4; Length 2885;
Best Local Similarity 60.1%; Pred. No. 6.7e-25;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 23 CTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAAGGGCTCTTAAACATCCAGAGA 82
Db 148 CTCAGAAGCTGACGAGCAGCACCAGCTCATCTTATCCAAAGCTTCTGAGGGAGGAGCA 207

QY 83 ACAGAGAAATGTGCTGACTGCAAAACAAAGGTCCAAAGTGGCTAGTGTATTATTAGTA 142
Db 208 ACAAGTACTGGCGGACTGCGAGGCCAAAGGTCTCCGATGGCTTCTCTGGAATATTGGTG 267

QY 143 TCTTATCTGCATGCATGTTCTGGGATTTACAGAGGTCTCGGGGTACACATATCGAAGG 202
Db 268 TGTATTATTTGCATCAGATGTGCTGGAATTCATAGAAATCTTGGGGTTTCAATATCCAGG 327

QY 203 TTCGATCTGCCACTCTGGACACATGCGCTCCCGAGCAGGTTGCATTTATACAGTCAATGG 262
Db 328 TCAATCAGTCAACTAGACCATGCGACAGCAGACAGATACAGTGCATGCAAGATATGG 387

QY 263 GAAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTACCCCCAAACTATGATAGA 318
Db 388 GAAATACTAAAGCAAGACTACTCTATGAAGCCAATCTTCCAGAGAACTTTCGAAGA 443

RESULT 4
US-09-484-970B-121
; Sequence 121, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 121
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 227709 (1221361CBI)
; US-09-484-970B-121

Query Match      6.5%; Score 94.2; DB 3; Length 1447;
Best Local Similarity 61.0%; Pred. No. 1.1e-20;
Matches 153; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 79 GAGAACAGAGAAATGTGCTGACTGCAAAACAAAGGTCCAAAGTGGCTAGTGTATTATTA 138
Db 110 GATAACAAGTTTGTGCGGATTGCCAGTCTAAAGGCCGCGATGGGCTCTTGGAAACATT 169

QY 139 GGTATCTTATCTGCATGCATGTTCTGGGATTCACAGGAGTCTCGGGGTACACATATCG 198
Db 170 GGTGTGTTTCATCTGCATTCGATGTGCTGGAATCCACAGGAATCTGGGGGTGCACATATCC 229

QY 199 AAGGTTGATCTGCCACTCTGGACACATGCGTCCCGAGCAGGTTGCATTTATACAGTCA 258
Db 230 AGGGTAAAGTCAAGTTAACTTCGACAGTGGACTCAAGAACAGATTCACTGATGCAAGAG 289

QY 259 ATGGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTACCCCCAAACTATGATAGA 318
Db 290 ATGGGAATGGAAGGCAACCGACTTTATGAAGCCTATCTTCTGAGACCTTTTCGGGGA 349

QY 319 GTTCGAATTGA 329
```

Db 350 CCTCAGATAGA 360

RESULT 5

US-09-620-312D-515
Sequence 515, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 515

LENGTH: 1976

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (47)...(1336)

US-09-620-312D-515

Query Match 6.5%; Score 94.2; DB 4; Length 1976;

Best Local Similarity 61.0%; Pred. No. 1.3e-20;

Matches 153; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 79 GAGAACAGAGATGCTGCTGACTGCAAAACAAAGGTCACAGATGGGCTAGTGTAAATTA 138

Db 116 GATAACAAAGTTTGTGCGAGATGCGGATGCTAAAGGCGCGGATGGGCTCTTGGAAACAT 175

Qy 139 GGTATCTTTATCTGCAATGTTCTGGATTCTCAGAGTCTCGGGGTACACATATCG 198

Db 176 GGTGTGTTTCATCTGCATTCGATGCTGGAATCCACAGGATCTGGGGGTGCACATATCC 235

Qy 199 AAGGTTTCGATCTGCCACTCTGGACACATGGCTCCCGAGCAGGTTGCATTTATACAGTCA 258

Db 236 AGGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 295

Qy 259 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 318

Db 296 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 355

Qy 319 GTTGGAAATGA 329

Db 356 CCTCAGATAGA 366

RESULT 6

US-09-513-999C-2719

Sequence 2719, Application US/09513999C

Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Duclert, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Patent No. 6783961

FILE REFERENCE: 59, US2, REG

CURRENT APPLICATION NUMBER: US/09/513,999C

CURRENT FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent.pm

SEQ ID NO 2719

LENGTH: 309

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 67..309

US-09-513-999C-2719

Query Match 5.9%; Score 86.2; DB 4; Length 309;

Best Local Similarity 59.0%; Pred. No. 1.8e-18;

Matches 148; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 79 GAGAACAGAGATGCTGCTGACTGCAAAACAAAGGTCACAGATGGGCTAGTGTAAATTA 138

Db 46 GCGGAAAGGAAACTGCTTAAATGATTTTAAAGGGCGGATGGGCTCTTGGAAACAT 105

Qy 139 GGTATCTTTATCTGCAATGTTCTGGGATTCTCAGAGTCTCGGGGTACACATATCG 198

Db 106 GGTGTGTTTCATCTGCATTCGATGCTGGAATCCACAGGATCTGGGGGTGCACATATCC 165

Qy 199 AAGGTTTCGATCTGCCACTCTGGACACATGGCTCCCGAGCAGGTTGCATTTATACAGTCA 258

Db 166 AGGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 225

Qy 259 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 318

Db 226 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 285

Qy 319 GTTGGAAATGA 329

Db 286 CCTCAGATAGA 296

RESULT 7

US-09-621-976-845

Sequence 845, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 845

LENGTH: 459

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 52..459

US-09-621-976-845

Query Match

5.9%; Score 86.2; DB 4; Length 459;

Best Local Similarity 59.0%; Pred. No. 2.4e-18;

	Matches	148;	Conservative	0;	Mismatches	103;	Indels	0;	Gaps	0;
Qy	79	GAGAACAGAGAAATGCTGACTGCAAAACAAAAGGTCCAAGATCGGCTAGTGTAAATTTA	138							
Db	31	CGCGAAAGGAAAACTGCTTAAATGATTTTAAAGGGCCGCGATGGGCTCTTGGAAACATT	90							
Qy	139	GGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGTCTCGGGGTACACATATCG	198							
Db	91	GGTGTGTTTCATCTGCATTCGATGTCTGGAATCCACAGGAATCTGGGGGTGCATATCC	150							
Qy	199	AAGTTTCGATCTGCCACTCTTGGAACACATGGCTCCCGAGCAGGTTGCAATTAACAGTCA	258							
Db	151	AGGGTAAAGTCAGTTAACTCTCGACCAGTGGACTCAAGAACAGATTCAGTGCATGCAAGAG	210							
Qy	259	ATGGGAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTACCCCAAACTATGATAGA	318							
Db	211	ATGGGAATTGAAAGGCNAACCGACTTTATGNAAGCCTATCTTCTGAGACCTTTCCGGGA	270							
Qy	319	GTTCGAATTTGA	329							
Db	271	CCTCAGATAGA	281							

```

RESULT 8
US-09-620-312D-1104
; Sequence 1104, Application US/09620312D
; Patent No. 6596662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6596662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1104
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (251)..(1300)
US-09-620-312D-1104

```

```

Qy 165 TGGGATTACAGAGAGTCTCGGGGTACACATATCGAAGTTTCGATCTGCCACTCTCGACAC 225
Db 166 TGGAAATCCACAGGAATCTGGGGGTGCACATATCCAGGGTAAAGTCAGTTAACTTCGACCA 225
Qy 225 ATGCTCCCCGAGCAGGTTCGATTATATACAGTCAATGGGAATGATAAAGCAAAATAGTTA 284
Db 226 GTGGACTCAAGAACAGATTTCAGTGCATGCAAGAGATGGGAATGGAAAGGCNAACCGACT 285
Qy 285 CTGGGAAGCAGAGCTACCCCCAAACTATGATAGAGTTGGAATTGA 329
Db 286 TTATGAAGCCTATCTTCTCGAGACCTTTCGGCGACCTCAGATGGA 330

RESULT 9
US-09-855-323-4
; Sequence 4, Application US/09855323
; Patent No. 6602667
; GENERAL INFORMATION:
; APPLICANT: Walker, Michael G.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: Inflammation-Associated Polynucleotides
; FILE REFERENCE: PB-0006-1 CIP
; CURRENT APPLICATION NUMBER: US/09/855,323
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1261
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 1221361CBI
US-09-855-323-4

Query Match 5.4%; Score 78.2; DB 4; Length 1261;
Best Local Similarity 60.7%; Pred. No. 2.6e-15;
Matches 128; Conservative 0; Mismatches 83; Indels 0; Gaps 0

Qy 119 GATCGGGCTAGTGTTAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGA 178
Db 51 GATCGGGCCTCTTGGAACTGGTGTGTTCATCTGCATGTCTGGAATCCACAGA 110
Qy 179 GTCTCGGGGTACACATATCGAAGTTTCGATCTGCACTCTGGACACATGGCTCCCGGAGC 238
Db 111 ATCTGGGGGTGCACATATCCAGGTAAGTCAGTTAACTTCGACCACTGAGTCAAGAAC 170
Qy 239 AGGTGTCATTATACAGTCAATGGGAATGATAAAGCAAAATAGTTATCTGGGAAGCAGAGC 298
Db 171 AGATTCACTGCATGCAAGAGATGGGAATGGAAAGGCAACCGACTTTATGAAGCCTATC 230
Qy 299 TACCCCCAAACTATGATAGAGTTGGAATTGA 329
Db 231 TTCTCGAGACCTTTTCGGCGACCTCAGATAGA 261

```

```

RESULT 10
US-09-854-133-660
; Sequence 660, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITL OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0

```

```
; SEQ ID NO 660
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-854-133-660

Query Match
  5.2%; Score 74.8; DB 4; Length 514;
Best Local Similarity 59.3%; Pred. No. 2e-14;
Matches 127; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 79 GAGACAGAGAAATGTCTGACTGCAAAACAAAGGTCCAGATGGCGTAGTGTAAATTTA 138
Db 283 GGGAACTCCCACTGTGTGACTGCGAGACCCAGAAATCCCAACTGGGCCAGTTTGAATCG 342
QY 139 GGTATCTTTATCTGATGCAATGTTCTGGATTTCACAGAGTCTCGGGGTACACATATCG 198
Db 343 GGAGCCCTCATGTGATCGAATGCTCAGGGATCCACCGGAATCTTGGCACCCACCTTTCC 402
QY 199 AAGGTTTCGATCGCCACTCTGCACACATGGCTCCCGAGCAGGTTTGCATTTATACAGTCA 258
Db 403 CGAGTCGGATCTCTGGACCTGGATGACTGGCCCAATCGAGCTCATCAAGTGATGTATCC 462
QY 259 ATGGGAATGATAAAGCAAAATAGTTACTGGGAAG 292
Db 463 ATCGGAACGAGCTAGCCCAACAGCGCTCTGGGAAG 496

RESULT 11
US-09-640-211A-1584
; Sequence 1584, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1584
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-09-640-211A-1584

Query Match
  5.0%; Score 72.2; DB 4; Length 435;
Best Local Similarity 58.7%; Pred. No. 1.3e-13;
Matches 125; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 74 ATCCGAGAACAGAGAAATGTCTGACTGCAAAACAAAGGTCCAGATGGCGTAGTGTTA 133
Db 144 AACCCGACAAAGATGCTGCTTGTGATGTATACAGGAATCCCATGCGGATCGGTCA 203
QY 134 ATTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGATCTCGGGGTACACA 193
Db 204 CTTACGGGATTTTCATCTGCTGGATGTTCTGCTATCTCATGCTAGTCTTGGTTCACA 263
QY 194 TATCGAGGTTTCATCTGCATCTGGACATGCTCCCGAGCAGGTTGCAATTTATAC 253
Db 264 TTAGCTTTTCTGATCTGTAACCTGGACTCTGGAATCTTGAACAGATTGAAGTCAATGA 323
QY 254 ACTCAATGGGAATGATAAGCAAAATAGTTACT 286
Db 324 GCTTTGGTGGCAATGGCCGAGGACATACATCTT 356

RESULT 12
US-09-620-312D-155
; Sequence 155, Application US/09620312D
```

```
; Patent No. 659662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 659662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: Pt FL_genes Version 1.0
; SEQ ID NO 155
; LENGTH: 6990
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(2397)
; US-09-620-312D-155

Query Match
  4.7%; Score 68.6; DB 4; Length 6990;
Best Local Similarity 58.6%; Pred. No. 1.7e-11;
Matches 119; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 91 TGTGCTGACTGCAAAACAAAGGTCCAGATGGCGTAGTGTAAATTTAGGTATCTTTATC 150
Db 1300 TGTGTGACTGTGGCTTGGCAGATCCACGGTGGCCAGCATCAACCTGGGCATCACCTTG 1359
QY 151 TGCATGCAATGTTCTGGGATTCACAGAGTCTCGGGGTACACATATCGAAGTTTCAATCT 210
Db 1360 TGTATCGAGTGTCTCGGAATTCACCGGAGCCTTGGGGTTCAATTTTCAAAAGTACGATCT 1419
QY 211 GCCACTTGGACACATGCTCCCGAGCAGGTTGCAATTTATACAGTCAATGGAATGAT 270
Db 1420 TTAATCTTTAGACACTCGGAGCCAGAACTTTTAAAGCTTATGTGTGAGTTGGGANTGAT 1479
QY 271 AAAGCAATAGTTACTTGGGAAGC 293
Db 1480 GTTATTAATCGAGTTTATGAAGC 1502

RESULT 13
US-09-513-959C-384
; Sequence 384, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
```

```
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 384
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 89..331
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: 89..208
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq LNLGVLMTECSG/IH
; US-09-513-999C-384

Query Match      4.7%; Score 68.4; DB 4; Length 332;
Best Local Similarity 57.5%; Pred. No. 2,1e-12;
Matches 123; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY   79  GAGACAGAGAATGTCTGACTGCACAAACAAGGTCCAAAGTGGCTAGTGTTAATTTA 138
     |||||
DB   119  GGGAATCCCACTCGCTGACTGTGAGACCAGCATCCCTAAGTGGGCCAGTTTGAAC TTG 178

QY   139  GGPATCTTTATCTGATGCAATGTTCTGGATTTCACAGAGTCTCGGGGTACACATATCG 198
     |||||
DB   179  GGAGTCTCTATGTTATGAATGTTTCAGGAATCCACCGCAGTCTTGGACCGCTTTCC 238

QY   199  AAGTTTCATCTGCCACTCTGACACATGGCTCCCCGAGCAGGTTCGATTATACAGTCA 258
     |||||
DB   239  CGGTGCGGATCTCTGGAGCTGAGTGGCCAGTTGAGCTCAGGAAGTTATGTCACT 298

QY   259  ATGGGAATGATAAACCAATAGTTACTGGGAG 292
     |||||
DB   299  ATTGCAATGACTAGCCAAACAGCATCCGGGAAG 332

RESULT 14
US-09-620-312D-35
; Sequence 35, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 15

; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (520)..(1614)
; US-09-620-312D-35

Query Match      4.5%; Score 65.4; DB 4; Length 2119;
Best Local Similarity 56.9%; Pred. No. 8.7e-11;
Matches 120; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY   82  AACAGAGAATGTCTGACTGCACAAACAAGGTCCAAAGTGGCTAGTGTTAATTAGGT 141
     |||||
DB   898  AACAGCTTTTGTATCGACTGGATGACCCAATCCAGACTGGCCAGCTGAACTGGGT 957

QY   142  ATCTTTATCTGCATGCAATGTTCTGGGATTTCACAGAGTCTCGGGGTACACATATCGAAG 201
     |||||
DB   958  GCCCTGATGTGCATTTGAGTGTCTCAGGCATCCACCGCACCTCGGGGGCTCACCTGTCCC 1017

QY   202  GTTCGATCTGCCACTCTGGACACATGCTCCCGACAGCTTGCATTTATACAGTCAA TG 261
     |||||
DB   1018  GTGCGCTCCCTTGACCTCGATGACTGCGCCCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 1077

QY   262  GGAATGATAAAGCAATAGTTACTGGGAAG 292
     |||||
DB   1078  GGCAATGCCCTCGCCACAGCGTCTGGGAGG 1108

RESULT 15
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZ9pt-Fls
US-08-232-463-14

Query Match      4.4%; Score 63.4; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 1e-09;
Matches 16; Conservative 228; Mismatches 149; Indels 0; Gaps 0;

Qy 264 AAATGATAAGCAATAGTTACTGGAGACGAGCTACCCCAAACTATGATAGAGTTGG 323
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1456 AAGAGATAGAGAAATTTGCTACRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1397

Qy 324 AATTGAGATTTTATACGTGCAAGTATCAAGAGAGAGATGGTTCTAGAGGGGAAA 383
Db :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1396 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1337

Qy 384 GGCTAGATCACCTCTAGAGTCAGCAGCAACGGCGGAAATCTGTGGAGAGAGTGGCC 443
Db :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1336 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1277

Qy 444 GGGATATGAGCATGGACATAGTAGTCTCTGTAATTTGTTGAGGAGAGGAAACTAT 503
Db :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1276 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1217

Qy 504 TCCAGCATAGACAAAGAAATAATCTTCTGCAACGAGATAAATCTCCCGTCCTCC 563
Db :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1216 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1157

Qy 564 CCAAGGACCCAGTCAGGTTATAAGCCACAGCAGAGAAATGGAGCTGTCAGCTACTCCAGT 623
Db :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1156 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1097

Qy 624 AGAGAGGGAGAGCAAGCAGTAAATGTTGCACC 656
Db :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1096 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1064
```

Search completed: September 27, 2005, 06:22:54
Job time : 285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 03:06:24 ; Search time 3611 Seconds
(without alignments)
2688.583 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaagcaacgt.....gaatgttcacaaacattga 1452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7400732 seqs, 3343137571 residues

Total number of hits satisfying chosen parameters: 14801464

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1452	100.0	1452	US-10-630-518-1	Sequence 1, Appli
2	1450.4	99.9	1452	US-10-630-518-3	Sequence 3, Appli
3	1450.4	99.9	1452	US-10-630-518-5	Sequence 5, Appli
4	432.6	29.8	2640	US-10-424-599-108715	Sequence 108715,
5	429.4	29.6	1879	US-10-425-114-10983	Sequence 10983, A
6	348.6	24.0	2154	US-10-767-701-14966	Sequence 14966, A
7	340	23.4	1951	US-10-425-114-27747	Sequence 27747, A

8	340	23.4	2299	20	US-10-425-115-176514	Sequence 176514,
9	338.4	23.3	2011	18	US-10-425-114-4099	Sequence 4099, Ap
10	338.4	23.3	2022	18	US-10-425-114-32446	Sequence 32446, A
11	336.6	23.2	1709	18	US-10-425-114-22038	Sequence 22038, A
12	336.6	23.2	1802	18	US-10-425-114-30303	Sequence 30303, A
13	336.6	23.2	1928	20	US-10-425-115-176512	Sequence 176512, A
14	335	23.1	1915	18	US-10-425-114-17368	Sequence 17368, A
15	333.4	23.0	2221	20	US-10-425-115-176513	Sequence 176513,
16	331.8	22.9	1902	18	US-10-425-114-2297	Sequence 2297, Ap
17	317.6	21.9	772	18	US-10-424-599-84194	Sequence 84194, A
18	307	21.1	2180	19	US-10-437-963-92697	Sequence 92697, A
19	298.2	20.5	596	19	US-10-021-323-3493	Sequence 3493, Ap
20	255.4	17.6	5580	19	US-10-630-518-8	Sequence 8, Appli
21	247.8	17.1	701	20	US-10-425-115-24352	Sequence 24352, A
22	227.4	15.7	313	9	US-09-294-093B-1317	Sequence 1317, Ap
23	227	15.6	779	20	US-10-425-115-176517	Sequence 176517,
24	213.6	14.7	2291	18	US-10-425-114-28016	Sequence 28016, A
25	213.6	14.7	2443	20	US-10-425-115-176518	Sequence 176518,
26	206.6	14.2	1407	20	US-10-739-930-3784	Sequence 3784, Ap
27	182.4	12.6	408	20	US-10-856-499-1383	Sequence 1383, Ap
28	155.8	10.7	528	18	US-10-424-599-105977	Sequence 105977,
29	155	10.7	732	20	US-10-425-115-176137	Sequence 176137,
30	147.8	10.2	598	20	US-10-425-115-176511	Sequence 176511,
31	127.2	8.8	290	9	US-09-294-093B-2723	Sequence 2723, Ap
32	122.4	8.4	365	18	US-10-424-599-136039	Sequence 136039,
33	116.8	8.0	388	20	US-10-425-115-87733	Sequence 87733, A
34	107.2	7.4	587	9	US-09-867-550-1487	Sequence 1487, Ap
35	107.2	7.4	2885	17	US-10-120-988-3	Sequence 1037, Ap
36	107.2	7.4	3194	21	US-10-887-553A-1037	Sequence 170957,
37	105.2	7.2	763	20	US-10-425-115-170957	Sequence 43, Appli
38	98.4	6.8	1403	17	US-10-426-776-43	Sequence 43, Appli
39	95	6.5	1084	18	US-10-425-114-31399	Sequence 31399, A
40	94.2	6.5	1976	15	US-10-037-270-515	Sequence 515, App
41	94.2	6.5	1976	17	US-10-117-722-515	Sequence 515, App
42	89.4	6.2	2263	20	US-10-425-115-39539	Sequence 39539, A
43	85.2	5.9	649	20	US-10-425-115-52793	Sequence 52793, A
44	82.6	5.7	1327	15	US-10-037-270-1104	Sequence 1104, Ap
45	82.6	5.7	1327	17	US-10-117-722-1104	Sequence 1104, Ap

ALIGNMENTS

RESULT 1
US-10-630-518-1
; Sequence 1, Application US/10630518
; Publication No. US20040143872A1
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: LILJEGREN, Sarah, J.
; APPLICANT: ECKER, Joseph, R.
; APPLICANT: YANOFKY, Martin, F.
; TITLE OF INVENTION: GENETIC CONTROL OF ORGAN ABSCISSION
; FILE REFERENCE: SALKINS.035C1
; CURRENT APPLICATION NUMBER: US/10/630.518
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: US 60/264,974
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCTUS02/01938
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1452)
US-10-630-518-1
Query Match 100.0%; Score 1452; DB 19; Length 1452;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1452; Conservative 0;

QY 1 ATGAACGAGAACCCACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
DB |||||
1 ATGAACGAGAACCCACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
QY |||||
DB |||||
61 GGGCTTTTAAACATCCAGAACACAGAGATGTGCTGATGCAACAAACAAAGGTCACAGA 120
DB |||||
61 GGGCTTTTAAACATCCAGAACACAGAGATGTGCTGATGCAACAAACAAAGGTCACAGA 120
QY |||||
DB |||||
121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGCATGCAATGTCTGGGATTCACAGGAGT 180
DB |||||
121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGCATGCAATGTCTGGGATTCACAGGAGT 180
QY |||||
DB |||||
181 CTCGGGGTACACATATCGAAGGTCGATCTGCCACTCTGCAACATATGCTGCCGAGCAG 240
DB |||||
181 CTCGGGGTACACATATCGAAGGTCGATCTGCCACTCTGCAACATATGCTGCCGAGCAG 240
QY |||||
DB |||||
241 GTTGCATTTATACAGTCAATGGAATGATAAGCAATAGTTACTTGGGAACGAGCTA 300
DB |||||
241 GTTGCATTTATACAGTCAATGGAATGATAAGCAATAGTTACTTGGGAACGAGCTA 300
QY |||||
DB |||||
301 CCCCCAACTATGATAGATTGGAATTTGAGAAATTTATACGTGCAAAATGATGAAGAG 360
DB |||||
301 CCCCCAACTATGATAGATTGGAATTTGAGAAATTTATACGTGCAAAATGATGAAGAG 360
QY |||||
DB |||||
361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCCGAGCAAGCGGG 420
DB |||||
361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCCGAGCAAGCGGG 420
QY |||||
DB |||||
421 AAATCTGTGGAGAGATGGCGCGGATATGAGCATGGACATAGTAGTCTGTAAAT 480
DB |||||
421 AAATCTGTGGAGAGATGGCGCGGATATGAGCATGGACATAGTAGTCTGTAAAT 480
QY |||||
DB |||||
481 TTGTTTGAGGAGGAGAAACTATTCAGCATCTAGAAACAAAGAAATATGCTGCAAG 540
DB |||||
481 TTGTTTGAGGAGGAGAAACTATTCAGCATCTAGAAACAAAGAAATATGCTGCAAG 540
QY |||||
DB |||||
541 AGAATAAATCTTCCGCTGCTCCCCAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
DB |||||
541 AGAATAAATCTTCCGCTGCTCCCCAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
QY |||||
DB |||||
601 ATGGAGTCTGACGCTACTCCAGTAGAGGGGAGAAACAGACAGTAAATGCTCCACGCA 660
DB |||||
601 ATGGAGTCTGACGCTACTCCAGTAGAGGGGAGAAACAGACAGTAAATGCTCCACGCA 660
QY |||||
DB |||||
661 TCAGATCTCTCCAAAGGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
DB |||||
661 TCAGATCTCTCCAAAGGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
QY |||||
DB |||||
721 TCGACTACAAATACCTCAGAGGCAATCTCTGGGATACCTCTGGCGATGATTAATCTATGG 780
DB |||||
721 TCGACTACAAATACCTCAGAGGCAATCTCTGGGATACCTCTGGCGATGATTAATCTATGG 780
QY |||||
DB |||||
781 GCTGGCTTTTCACTGCTGGAAGTGTGCTCAACGGCAGAGAAATGCTACAGCCAGCCT 840
DB |||||
781 GCTGGCTTTTCACTGCTGGAAGTGTGCTCAACGGCAGAGAAATGCTACAGCCAGCCT 840
QY |||||
DB |||||
841 GCTGAGCAGATGCTCTCTCAGCTTCAATCTTCTGACTTTGAGATTTGTTTAAAGCACA 900
DB |||||
841 GCTGAGCAGATGCTCTCTCAGCTTCAATCTTCTGACTTTGAGATTTGTTTAAAGCACA 900
QY |||||
DB |||||
901 CCTAATTTAACTCAACCAAGCAACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
DB |||||
901 CCTAATTTAACTCAACCAAGCAACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
QY |||||
DB |||||
961 GAGAAGCAATATAGTATGCGCTTTTGGCATGCAATCAGCAACAGGTTCTATGCTCGCT 1020
DB |||||
961 GAGAAGCAATATAGTATGCGCTTTTGGCATGCAATCAGCAACAGGTTCTATGCTCGCT 1020
QY |||||
DB |||||
1021 CAGCAGAACGCTTTTACATGCTGAGGCAAGGCTGCTGGAGGCACTCCAAACGGCGTG 1080
DB |||||

RESULT 2

US-10-630-518-3
; Sequence 3, Application US/10630518
; Publication No. US20040143872A1
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: LILJEGREN, Sarah, J.
; APPLICANT: ECKER, Joseph, R.
; APPLICANT: YANOFSKY, Martin, F.
; TITLE OF INVENTION: GENETIC CONTROL OF ORGAN ABSCISSION
; FILE REFERENCE: SALKINS.035C1
; CURRENT APPLICATION NUMBER: US/10/630,518
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: US 60/264,974
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCTUS02/01938
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1452)
US-10-630-518-3

Query Match 99.9%; Score 1450.4; DB 19; Length 1452;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAAGCAAGCTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
DB |||||
1 ATGAACGAGAAAGCAAGCTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
QY |||||
DB |||||
61 GGGCTTTTAAACATCCAGAACACAGAGATGTGCTGATCCTCAAAACAAAGGTCACAGA 120
DB |||||
61 GGGCTTTTAAACATCCAGAACACAGAGATGTGCTGATCCTCAAAACAAAGGTCACAGA 120
QY |||||
DB |||||
121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCAATGTTCTGGAATTCACAGGAGT 180
DB |||||

121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTACATGCAATGTTCTGGGATTTCACAGGAT 180
181 CTGGGGTACATATCGAAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
181 CTGGGGTACATATCGAAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
241 GTTGTCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
241 GTTGTCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
301 CCCCCAACTATGATAGAGTTGGAATTTAGCAATTTTATACGTCCAAAGTATGAAGAAG 360
301 CCCCCAACTATGATAGAGTTGGAATTTAGCAATTTTATACGTCCAAAGTATGAAGAAG 360
361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTTAGAGTCGAGCAGGACCGCG 420
361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTTAGAGTCGAGCAGGACCGCG 420
421 AAATCTGTGAGAGAGTGGCGCGGATATGAGCATGGACATAGTAGTCTCTGTAAT 480
421 AAATCTGTGAGAGAGTGGCGCGGATATGAGCATGGACATAGTAGTCTCTGTAAT 480
481 TTGTTTGGAGAGGAAAACCTATTTCCAGCATCTAGAACAAAGAAATATGTTGCTGCAACG 540
481 TTGTTTGGAGAGGAAAACCTATTTCCAGCATCTAGAACAAAGAAATATGTTGCTGCAACG 540
541 AGAATAAATCTCCGTCCTCCCAAGGACCCAGTCAGTCAAGTAAATGTTGACAGCAAGAA 600
541 AGAATAAATCTCCGTCCTCCCAAGGACCCAGTCAGTCAAGTAAATGTTGACAGCAAGAA 600
601 ATGAGTCTCAGCTACTCCAGTAGAGAGGAGGAGAAACAAGCAGTAAATGTTGACAGCA 660
601 ATGAGTCTCAGCTACTCCAGTAGAGAGGAGGAGAAACAAGCAGTAAATGTTGACAGCA 660
661 TCAGATCTCCCAAGGTTGATTTGCTACTGATCTGTTTAACTGCTCAATGGAAT 720
661 TCAGATCTCCCAAGGTTGATTTGCTACTGATCTGTTTAACTGCTCAATGGAAT 720
721 TCAGTACAAATACCTCAGAGGCACTCTCCGCGATCTCTCCGCGATGATCACTATG 780
721 TCAGTACAAATACCTCAGAGGCACTCTCCGCGATCTCTCCGCGATGATCACTATG 780
781 GCTGGCTTTTCAGTCTGCTGAAGTGTCAACCGCAGAGAAATGTCAAGCAAGCCT 840
781 GCTGGCTTTTCAGTCTGCTGAAGTGTCAACCGCAGAGAAATGTCAAGCAAGCCT 840
841 GCTGAGAGCAGTTCTCTCCAGCTTCACTCTCTGACTTTGAGGATTTGTTTAAAGCAC 900
841 GCTGAGAGCAGTTCTCTCCAGCTTCACTCTCTGACTTTGAGGATTTGTTTAAAGCAC 900
901 CCTAATTTAACTCAAGCACCAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
901 CCTAATTTAACTCAAGCACCAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
961 GAGAGAGCAATATAGTATCGCTTTTGCCATCATCAGCAACAGGTTGCTATGCTCGCT 1020
961 GAGAGAGCAATATAGTATCGCTTTTGCCATCATCAGCAACAGGTTGCTATGCTCGCT 1020
1021 CAGCAGCAAGCCTTTTACATGGCTGCGAGCAAGAGCTGCTGGAGGCACTCCAAACGGCGTG 1080
1021 CAGCAGCAAGCCTTTTACATGGCTGCGAGCAAGAGCTGCTGGAGGCACTCCAAACGGCGTG 1080
1081 AATCAACAGCTATGCTAATGCTCTTAAAGTGTCTGCAAAATGGTCAAAACCCCGG 1140
1081 AATCAACAGCTATGCTAATGCTCTTAAAGTGTCTGCAAAATGGTCAAAACCCCGG 1140
1141 GGCTACAGATCCCGGAATGCTAATCCCGGTAGGTGCTCAAGCTGATCTCCAGAACTT 1200
1141 GGCTACAGATCCCGGAATGCTAATCCCGGTAGGTGCTCAAGCTGATCTCCAGAACTT 1200
1201 ATGCAAAACATGATGAACGCAACATGAACGAGACCCGCAACCAAGAGAGAAC 1260

1201 ATGCAAAACATGATATGAACGCAACATGAACACGAGACCCGCAACCAAGAGAAC 1260
1261 ACTCTACATATCCCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
1261 ACTCTACATATCCCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
1321 ATGACCCCAAACTCAACCGGTAAACCTCAGTCATCATCCGCAACCAACCAAGCACC 1380
1321 ATGACCCCAAACTCAACCGGTAAACCTCAGTCATCATCCGCAACCAACCAAGCACC 1380
1381 ACACATCTTCAATCAGGCAAGACTTTGATTTCTCTCTTGTGATGATGGAATGTTT 1440
1381 ACACATCTTCAATCAGGCAAGACTTTGATTTCTCTCTTGTGATGATGGAATGTTT 1440
1441 ACAAACATTTGA 1452
1441 ACAAACATTTGA 1452

RESULT 3
US-10-630-518-5
; Sequence 5, Application US/10630518
; Publication No. US20040143872A1
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: LILJEGREN, Sarah, J.
; APPLICANT: ECKER, Joseph, R.
; APPLICANT: YANOFKY, Martin, F.
; FILE OF INVENTION: GENETIC CONTROL OF ORGAN ABSCISSION
; FILE REFERENCE: SALKINS.035C1
; CURRENT APPLICATION NUMBER: US/10/630,518
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: US 60/264,974
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCTUS02/01938
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(594)
US-10-630-518-5

Query Match 99.9%; Score 1450.4; DB 19; Length 1452;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCATAGAAAGATTCTTGAA 60
1 ATGACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCATAGAAAGATTCTTGAA 60
61 GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTCCAAGA 120
61 GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTCCAAGA 120
121 TGGGCTAGTGTGTTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTTCAAGAGT 180
121 TGGGCTAGTGTGTTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTTCAAGAGT 180
181 CTCGGGGTACATATCGAAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
181 CTCGGGGTACATATCGAAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
241 GTTGTCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
241 GTTGTCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
301 CCCCCAACTATGATAGAGTTGGAATTTAGCAATTTTATACGTCCAAAGTATGAAGAAG 360

1381 ACACATCTTACAAATCAGCAAGACCTTGTATTTCTCTTCTGATGATGGAATGTTTC 1440

1441 ACATAACATTGA 1452

1441 ACATAACATTGA 1452

RESULT 4

US-10-424-599-108715

Sequence 108715, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 108715

LENGTH: 2640

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MBT3847_69183C.1

US-10-424-599-108715

Query Match 29.8%; Score 432.6; DB 18; Length 2640;

Best Local Similarity 61.5%; Pred. No. 3.2e-122;

Matches 914; Conservative 0; Mismatches 499; Indels 72; Gaps 11;

1 ATGACGAGAAAGCCCAAGCTCTCTAAGGAGCTTATGCGCCATAGAAAGATTTCTTGAA 60

215 ATGAAACAGCAAGGCGCAACGTTTCCAAAGAGCTCAACGCCAAGCAAGANGATCTGGAA 274

61 GGGCTCTTTAAACATCCAGAGAAACAGAGAAATGTGTGACTGCAAAACAAAGAGTCCAAGA 120

275 GGACTTTCTTAATATACAGAGAAATAGGGAATGTGCTGACTGCAAGCTAAAGTCCAAGA 334

121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGATGCAATTTCTTGGGATTCACAGGAT 180

335 TGGGCTAGTGTAAATTTAGGTATCTTTATGATGCAATTTCTTGGGATTCACAGGAT 394

181 CTCGGGTACACATATCAAGGTTCCGATCTGCACTGCAACATGCTGCAACATGCTGCAAC 240

395 TTGGGGGTACATATATCAAGGTTCTGTTCTGCAACCTTGGACACTTGGCTTCCAGAGCAA 454

241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTTACTTGGGAAGCAGACTA 300

455 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTTACTTGGGAAGCAGACTA 514

301 CCCCCAAATATATGATAGAGTTGGAATTTGAGAAATTTTACGTGCAAGTATGAGAGAGAG 360

515 CTTCCAAATATATGATAGAGTTGGAATTTGAGAAATTTTACGTGCAAGTATGAGAGAGAG 574

361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCTGAGCAGGAAACGGCG 420

575 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCTGAGCAGGAAACGGCG 631

421 AAATCTGTGGAGAGAGAGTGGGCGGGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 480

632 AGTCTTTCACATTTGCGAGAGGCTGTGGAGAGAGTGGTTATGCGCTGTTTCTGAAAT 691

481 TTGTTTGGAGAGAGAGAGAGTATTTCCAGCATCTAGAACACAGAAATATATGTTGCTGCAAC 540

692 AAATTTGAGAGAGAGAGAGAGTATTTCCAGCATCTAGAACACAGAAATATATGTTGCTGCAAC 751

541 AGAATAAATCTTCCGCTGCTCCCAAGGACCCAGGCTAGGTTATATAAGCCACAGC----- 595

752 GTTCTGCTCTCCCAAGGAGTCTGAGCAGGTAATCTCTATTTACCAAACTCAGCATGTG 811

301 CCCCCAACTATGATAGAGTTGGAATTTAGAAATTTTATACGTGCAAAATATGAAGAGAG 360

361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCTGAGCAGAGAGAGCGCG 420

361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCTGAGCAGAGAGAGCGCG 420

421 AAATCTGTGAGAGAGAGTGGGCGGGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 480

421 AAATCTGTGAGAGAGAGTGGGCGGGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 480

481 TTGTTTGGAGAGAGAGAGAGTATTTCCAGCATCTAGAACACAGAAATATATGTTGCTGCAAC 540

481 TTGTTTGGAGAGAGAGAGAGTATTTCCAGCATCTAGAACACAGAAATATATGTTGCTGCAAC 540

541 AGAATAAATCTTCCGCTGCTCCCAAGGACCCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 600

541 AGAATAAATCTTCCGCTGCTCCCAAGGACCCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 600

601 ATGGAGTCTGAGTACTCTCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

601 ATGGAGTCTGAGTACTCTCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

661 TCAGATCTCTCAAAGTGGATTTTGTCTGATCTGTCTGATCTGTCTGATCTGTCTGATCTGT 720

661 TCAGATCTCTCAAAGTGGATTTTGTCTGATCTGTCTGATCTGTCTGATCTGTCTGATCTGT 720

721 TCAGTCAATAATCTCAGAGGCACTCTCTGGGAGTACTCTCTGGGAGTACTCTCTGGGAGT 780

721 TCAGTCAATAATCTCAGAGGCACTCTCTGGGAGTACTCTCTGGGAGTACTCTCTGGGAGT 780

781 GCTGGCTTTTCAAGTGGTGTCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

781 GCTGGCTTTTCAAGTGGTGTCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

841 GCTGAGAGAGTCTCTCTCAGGCTTCTCTGATCTGATCTGATCTGATCTGATCTGATCTG 900

841 GCTGAGAGAGTCTCTCTCAGGCTTCTCTGATCTGATCTGATCTGATCTGATCTGATCTG 900

901 CCTAATTTAACTCAACAGCACCACCAAGAGTGTGAAGGCGCATATCATGAGCCTGTTT 960

901 CCTAATTTAACTCAACAGCACCACCAAGAGTGTGAAGGCGCATATCATGAGCCTGTTT 960

961 GAGAAGCAATATAGTATCGCTTTTGGCATGATGATGATGATGATGATGATGATGATGATG 1020

961 GAGAAGCAATATAGTATCGCTTTTGGCATGATGATGATGATGATGATGATGATGATGATG 1020

1021 CAGCAGCAAGCCTTTTACATGGCTGCGAGGAAAGCTGCTGGAGGCACTCCAAACCGCGTG 1080

1021 CAGCAGCAAGCCTTTTACATGGCTGCGAGGAAAGCTGCTGGAGGCACTCCAAACCGCGTG 1080

1081 AATCAACAGCTATGCTATGCTCTTACGTAGCTTCTGCAATTTGGTCAAAACCCCGCG 1140

1081 AATCAACAGCTATGCTATGCTCTTACGTAGCTTCTGCAATTTGGTCAAAACCCCGCG 1140

1141 GGCTACAGATCCCGGATGATCAACCCGATGATGATGATGATGATGATGATGATGATGATG 1200

1141 GGCTACAGATCCCGGATGATCAACCCGATGATGATGATGATGATGATGATGATGATGATG 1200

1201 ATGCAAAACATGATATGAG 1260

1201 ATGCAAAACATGATATGAG 1260

1261 ACTCTCAATACCCATCATCTCTTACAGTGGTCAAGTGGTCAAGTGGTCAAGTGGTCAAGT 1320

1261 ACTCTCAATACCCATCATCTCTTACAGTGGTCAAGTGGTCAAGTGGTCAAGTGGTCAAGT 1320

1321 ATGACCCCAACTCAACCGGTAACCTCAGTCAATCATCGCAACCCCAACCAACCAACCAAC 1380

1321 ATGACCCCAACTCAACCGGTAACCTCAGTCAATCATCGCAACCCCAACCAACCAACCAAC 1380

1381 ACACCATTTCAAACTCAGGCAAGAGTTGATTTCTCTCTTGTGATGATGATGATGATGATG 1440

```

, TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
, TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
, FILE REFERENCE: 38-21(53313)B
, CURRENT APPLICATION NUMBER: US/10/425,114
, CURRENT FILING DATE: 2003-04-28
, NUMBER OF SEQ ID NOS: 73128
, SEQ ID NO 10983
, LENGTH: 1879
, TYPE: DNA
, ORGANISM: Glycine max
, FEATURE:
, OTHER INFORMATION: Clone ID: 700994349_FLI
US-10-425-114-10983

```

Query Match	29.6%	Score 429.4	DB 18	Length 1879
Best Local Similarity	61.4%	Pred. No. 2.5e-121		
Matches	912	Conservative	0	Mismatches 501; Indels 72; Gaps 11;
Qy	1	ATGAACGAGAAAGCCCAACGCTCTCTAAGGAGCTTAATGCGCGCCATAGAAGATTCCTTGAA	60	
Db	198	ATGAACGAGAAAGCCCAACGCTTTCAAAGAGCTCAAAGCCCAAGCACAAGAAGATACTCGAA	257	
Qy	61	GGCGTCTTTAAACATCCAGAGAAACAGAGAATGTGCTGACTGCAAAACAAAGGTCGAAGA	120	
Db	258	GGACTTCTTAAATFACCAGAGATAGGGAATGTGCTGACTGCAAAAGCTAAGGTCGAAGA	317	
Qy	121	TGGGCTAGTGTTAAATTTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT	180	
Db	318	TGGGCTAGTGTTAAATCTTGGCATCTTTATATGATGACGAGTGTTCAGGAATACATCTCGAAGT	377	
Qy	181	CTCGGGGTACACATATCGAAGGTTTCGATCTGCCACTCTGACACATGGCTCCCGAGCAG	240	
Db	378	TTGGGGGTACATATATCAAAGGTTCTGTTCTGCAACCTTGACACTTGGCTTCAGAGCAA	437	
Qy	241	GTTCGATTTATACAGTCAATTCGGAATATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA	300	
Db	438	GTTCGATTCATTCAAATCGGAACGAGAAAGCAAAATGTTTCTGGGAAGCAGAAATTA	497	
Qy	301	CCCCAAACTATGATAGAGTTGGAAATGAGAAATTTATACGTGCAAAAGTATGAAGAGAAG	360	
Db	498	CCTCCAAATTTATGATAGAGTTGGAAATGAGAAATTTCAATCTGTCAAAGTATGATGAAGAAG	557	
Qy	361	AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTCTAGAGTCGAGCAGGAAACGCGG	420	
Db	558	AGATGGGTTTCCAAGAGATGGGAAATTCAAAACACCTTCTCGAAT---TCGGGAAGAGAAA	614	
Qy	421	AAATCTCTGAGAGAGAGTGGCGCGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT	480	
Db	615	AGTCCTTCAATGGCAGAGCGCTGTGGAGAGAGTGGTTATCGCGCTGTTTCTGAAAT	674	
Qy	481	TTGTTTGAGGAGAGGAAAACATAATCCAGCATCTAGAAACAGAAATAATGTTGCTGCAACG	540	
Db	675	AAATTTGAGGAAAGGAAAGAAAATCCAAACCATCAACTGCAATTCCTGCGCGCAAGAAATTAAT	734	
Qy	541	AGAAATAATCTTCCCGTGCCTCCCAAGGACCCAGTCAGGTTATAAGCCACACAGC-----	595	
Db	735	GTTCCTCTCTCTCCAGAGCACCTGACGAGTAATCTCTATACCAACCTCAGCATGTG	794	
Qy	596	-AGAAATGAGTCTGCAGCTACTCTCCAGTAGAGAGGGAGAAAC-----AAGCAGTA	645	
Db	795	GAGAAATGGAATCAGTAGACCAACCAACCAACCAAGCTCCCCAACCATAGCTGAAACATCA	854	
Qy	646	AATGTTGCACCAAGCATCAG-----ATCCTCCAAAGGTGGATTTTGCTACTGATCTG	696	
Db	855	AAGCAGGCTACAGACACAGTTTAAATAATACCCCTCTAAAGTTGACTATGCCACAGACCTT	914	
Qy	697	TTTAAACATGCTATCAATGGATGATTCGACTACAAATACCTCAGAGGCAATCTCCTGGCGAT	756	
Db	915	TTCAACATGTGTCTATGGATGGGCCCAATGAAATGGCTCTGAGGCAGCT-----GGT	968	
Qy	757	ACTCCTGCCGATGATACTCATGGCTGGCTTTTCAGTCTGCTGGAAGTGGTCAAAACGGCA	816	
Db	969	ACAACTACTGATGATAATCACTGGGCGAGGTTTCCAGTCTGCTGCAGAGGTGTCAAACGCT	1028	

Qy	596	-AGAAAATGGAGTCTG	CAGCTACTCTC	AGTAGAGGGAGAAAC-----AAGCAGTA	645
Db	812	GAGAAAGTGGAAAC	CAGTAGACCA	CAACCAAGCTCCCCAACCATCTAGCTGAAACATCA	871
Qy	646	AATGTTGCACCA	CAGCATCAGAT-----CCTCCAAAGTGGATTTTGCTACTGATCTG	696	
Db	872	AAGCAAGCTAC	GAGACACAGTTCAAAATATCCCTCCAAAGTTGACTATGCTACAGACCTT	931	
Qy	697	TTTAAACATGCTAT	CAATGGGATGATTCGACTACAAATACCTCAGAGGCAACTCCTGGCGCAT	756	
Db	932	TTCAACATGTTGT	CTATGGATGGCCCTAATGMAAATGGCTCTGAGGCAGCT-----GGT	985	
Qy	757	ACTCCTGCCGAT	ATACTATNGGGCTGGCTTTTCAGTCTGCTGGAAGTGTGTCAAAAGCA	816	
Db	986	ACAACACTATGAT	ATACTCACTGGGCAGGTTTTTCAGTCTGCTGCAGAGGTGTCAACAGCT	1045	
Qy	817	GAGAAATTTGTC	CAGCCAGCCCTGCTGAGAGCAGTTCTCCTCCAGCTTCATCTCTGA-	875	
Db	1046	GAGAAGACTAGT	CCCTCCGAAAGCAGCTGATAGTACTCCAGGTTCTGCACTGGGAATTGAG	1105	
Qy	876	-----CTTTGAGGATTTGTTTAA	GGAACAACACTTAATTTAAACCTCAACAAAGCACCAGAAA	930	
Db	1106	GATCTTTTCAAGAATTTACATCCTG	TGACACCAAGTTTGACTCCAGAAATAACACAGAA	1165	
Qy	931	GATGTGAAGCGATATCATGAGCCTG	TTTGAGAAGACGAATATAGTATCGCTTTTGCC	990	
Db	1166	GATGTGAAAATGATATCATGAGCCTCTTTG	AGAAGGCAATATGGTGTCTCCATTTTCT	1225	
Qy	991	ATGCATCAGAAACAGTTGCTATGCTCGCTC	AGCAGCAAGCCCTTTACATGG---CTGCA	1047	
Db	1226	ATGCATCAACAGCAGCTTGGCCATGCTAGCA	CAGCAACAGTCTCTCTTAATGGGCTTTGCA	1285	
Qy	1048	GCGAAGCTGCTGAGGCACTCCAAA	CGGCGTGAATCAAAGCTATTTGCTTAATGCTCTT	1107	
Db	1286	GCTAAATCTTCTGGTGGGATCTCTA-----	GGTATCTGTGTAGCATACACACCTAGA	1339	
Qy	1108	AACGTAGCTTCTGCAAAATTTGGTCAAA	CCCCGGCGGTACCAGATCCCGGAATGACTAAC	1167	
Db	1340	CCCAATGTTCTATCCAAAGTTGCCAGCTACTG	CTATCAATTCCTGTGGAGT---AATG	1396	
Qy	1168	CCCGTAGTGTGTAAGTGTATCTCAGAAA	CTTATGCAAAACATGAATATGAAGCMAAC	1227	
Db	1397	CCCATGGTGTCTAGGAGAGAGTTGCAAAA	ACTTGTGAGGCTCAGACT-----1444		
Qy	1228	ATGAACAGAGACCGGACACACCCAGAGACACTCT	TACAATACCCTCATCCAGTTTC	1287	
Db	1445	AGGAACATGACCCGACGACATTTTGACGGGAGT	TCTGTCTCAATATCACCATCCAGTTTC	1504	
Qy	1288	TACAAATGGGTCAAGCTTAATCAAGTGAACGGT	ATGACCCCAACTCAAACGGTAAACCT	1347	
Db	1505	TATGGTATGGGACAAAGTTCTCTCCAGTTAAC	CGGTATGACAAATGGGAGTGAAGTAAACCT	1564	
Qy	1348	CAGTCATCATCCGCAACCCCAACCAAGCA	CCACCAATCTTCACAAATCAGGCAAGAC	1407	
Db	1565	CAGTCAGGAGCTCTCGCAGTGTCTATCTACCA	-----CTTCACAGTCTCGAAGGAT	1615	
Qy	1408	TTTGATTTTCTCTTCTCTGATGGATGTTTCA	CAAAACATTGA	1452	
Db	1616	TATGATTTTTTCTCTCTTAAACAGGGAATG	TTTATGAACAGTGA	1660	

```

RESULT 5
US-10-425-114-10983
; Sequence 10983, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Taskaka, Jack E
; APPLICANT: Cao, Yongwei

```

817 GAGAAATGTCACAGCCAGCTGCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGAC 876
1029 GAGAGAGCTAGTCCCTCGAAGCAGCTGATAGTACTCCAGGTTCTGCACTCTGGAATTGAG 1088
877 -----TTTGAGGATTTGTTTAAAGGACACACCTTAATTTAAACAACTCAACAAGCAACAAAA 930
1089 GATCTTTTCAAAAGATTTACATCTGTGACACCAAGTTTGTACTCCAGAAAAACACAGAAA 1148
931 GATGTGAAGGGGATATCATGAGCCCTGTTTGAGAGACCAATATAGTATCGCCTTTTGGC 990
1149 GATGTGAAGAAATGATATCATGAGCCCTTTTGAGAGGGCAATATGTTGTTCTCAATTTCT 1208
991 ATGCATCAGCAACAGAGTTGCTATGCTGCTCAGCAGCAAGCCCTTTTACATGG---CTGCA 1047
1209 ATGCATCAGCAGCTTGGCCATGCTAGCAGCAGCAAGCTCTCTTCTAATGCTCTTCTGCA 1268
1048 GGAAGAGCTGCTGGAGGCACTCAACAGCGGTGAATCAACAAGCTATTGCTTAATGCTCTT 1107
1269 GCTAAATCTTCTGTTGGGATCCTA-----GGTATCTCTGTAGCATACAACAACCTAGA 1322
1108 AACGTAGCTTCTGCAAAATTTGTCAAACCCGGGGCTACAGATCCCGGAATGACTAAC 1167
1323 CCCAATGTTCTTATCCAAAGTTGGCCAGCTACTGGCTATTCAATTCCTGGAGT---AATG 1379
1168 CCCGTAGTGGTCAAGCTGATCTCCAGAAATTTATGCAAAACATGAATATGAACGCAAAAC 1227
1380 CCCATGGTGGTTCAGGAGAGTTGCAAAACTTTGTGAGGCTCAGACT----- 1427
1228 ATGAACACAGAGCCCGCAACACCGGAGAGAACACTTCAATATACCATCATCAGTTTC 1287
1428 AGGAACATGACCCAGCACAATTTGAGGGAGTTCTGTTCATATATCCACCATCCAGTTTC 1487
1288 TACACAACTGGTCAAGCTAATCAAGTGAAGCTATGACCCCAAACTCAACCGTTAACT 1347
1488 TATGTTATGGGCAAGTTCTCTCTCAGTTAAAGGTATGACCAATGGGAGTGAATAACT 1547
1348 CAGTCATCATCCGCAACCCCAACCAAGCAGCACCACCATCTTCAATATCAGGCAAGAAC 1407
1548 CAGTCAGGAGCTCTCCTCAGTTCACTACCA-----CTTCAGCTCTCGAAGGAT 1598
1408 TTTGATTTCTTCTCTTGTGATGGAATGTTTCAACAAACATGTA 1452
1599 TATGATTTTCTCTTAAACAGGGAATGTTTATGAACAGTGA 1643

RESULT 6

US-10-767-701-14966
; Sequence 14966, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 14966
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS849_1
US-10-767-701-14966

Query Match 24.0%; Score 348.6; DB 19; Length 2154;
Best Local Similarity 60.3%; Pred. No. 2.8e-96;
Matches 645; Conservative 0; Mismatches 394; Indels 30; Gaps 3;
1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCGCCCATAGAAAGATTCTTGAA 60

RESULT 7

251 ATGAACGAGAAAGCGTCCGTTTCAAGGAGCTCAAGCCCAAGCAACAAGATATTGAA 310
61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGCTGACTGCTCAAAACAAGAGTCCAGA 120
311 AGTCTTCTTCCGCTTCTGAGAAATAGAGAAATGTCGGACTGCAAGTCAAGGGTCTCGA 370
121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCAATGTTCTTGGATTCACAGGAT 180
371 TGGCAAGTGTAAATCTAGGGATCTTTATATGATGATCATGTTCTTGGAAATTCATAGAAGC 430
181 CTCGGGGTACACATATCGAAGGTTGATCTGCACTCTGGACACATGCTCCCGAGCAG 240
431 CTGGGGGTGACATATCTAAGGTAAAGATCTGCAACCCCTGGATACATGCTCCCGAGCAA 490
241 GTTCATTTATACAGTCAATGGGAAATGATAAAGCAATAGTTACTTGGGAAAGCAGAGCTA 300
491 GTTGCAATTTATCAATCAATGGGAAACGAAAGCAATAGCTATTGGGAAGCAGAGCTG 550
301 CCCCAAACTATGATAGAGTTGGAATTTGAGAAATTTATACGTGCAAAAGTATGAAGAGAG 360
551 CCTCTAATATGATAGGTTGGGATAGAGAAATTTCACTCCGTGCAAAATATAGAGACAAG 610
361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCGACAGCAAGCGGG 420
611 AGATGGGTACCCGAGGAATGGAAACATCAGACCTACCTCCGGTGTTCGAGATGAGAAGAGC 670
421 AAATCTGTGGAGAGAGTGGSCCGGATATGAGCATGAGCATAGTAGTAGTCTCTTAAT 480
671 CAAGAGTCTCTGGCTAGTGTCTAATAGAGTGAACATGCTCATAGATCTTCAATTTGAGCAA 730
481 TTGTTTGGAGGAGAGAAACTTATTCAGCATCTAGAACCAAGAAATAATTTGCTGCAACG 540
731 AACCGTTTACAGCTCTTCCGAGCAAGTTGCACTCTAGCTTCAAGGATACCTCTCTCAG 790
541 AGAATAATCTTCCCGTCCCTCCCAAGGACCCAGTCAAGTTTAAAGCCACAGCAGAAA 600
791 GCATCACCTCAGCCGCCCAAGGTTAGAACCAAGTCTTAAGGTGGTTTCACTCTCTCAG 850
601 ATGAGTCTGCACTCTCAGTAGAGAGGAGGAGAAACAAGCAGTAAATTTGTCACACAGCA 660
851 CCACAGAAATCTCTGCAAAAGTTGAGGCAACACCCCTTAAAGTTGAGAAGCCATCAGTT 910
661 TCAGATCTTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAACTGCTATCAATGATGAT 720
911 GCGCGCTCTCTAAAGTCGATTATGCTACCGATCTCTTAAACATGTTATCAATGGATGGA 970
721 TCGACTACAATACCTCAGAGGCAACTCTCTGGCGATACTCTCTGCGATGATAACTCATGG 780
971 ACAACAGAGAAAGAGTCAGAGTCTATCTTCAAAT-----GATGATAATGCCCTGG 1018
781 GCTGGCTTTTCACTCTGCTGGAAGTGGTCAAAACGGCAGAGAAAAATTTGTCAAGCCCAAGCT 840
1019 GATGGATTCAGTCTGCAACACCAAGTACCTAGCTCAGAGAAAAAAGATTTCTGCCAAACG 1078
841 GCTGAGAGCAGTCTCTCTCCAGTTCATCTTCTGACTGTTGAGGATTTGTTTAAAGACACA 900
1079 GCAGAAAGCAAGACCC-----AGTCAACATCCGGAATAGAAGACCTATTTTAAAGACTCA 1132
901 CCTAATTTAACAACCTCAACAGCACC-----AAAAGATGTCAAAAGCGGATATC 948
1133 CCAGCTCTGTCAATATCTCTCAGCTCCAGCTGTTTCCCAAGTAAATGCGAAGATGATATC 1192
949 ATGAGCCTGTTTGAAGAGACGAATATAGTATCGCTTTTGGCCATGCTATCAGCAACAGGTT 1008
1193 ATGAGTTTGTGAGAGTCCATATGTTATGCTATCGCTTTGCTGCCCATCAACAGCAGCTG 1252
1009 GCTATGCTCGCTCAGCAGCAAGCCCTTTTACATGCTGCGAGGAAAGCTG 1057
1253 GCGTTTATGCTCAGCAGCAAGCTCTTCTAATGCTGCTCTTAAAGCTG 1301

US-10-425-114-27747
; Sequence 27747, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27747
; LENGTH: 1951
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4739-018-P5_FLI
US-10-425-114-27747
Query Match 23.4%; Score 340; DB 18; Length 1951;
Best Local Similarity 61.0%; Pred. No. 1.2e-93;
Matches 654; Conservative 0; Mismatches 385; Indels 33; Gaps 5;
QY 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCCGCATAGAAAGATTCTTGAA 60
DB 259 ATGAACGAGAAAGCGTCCGCTCTCCAGGAGCTCAAGCCAGCACAGAAAGATATTGAA 318
QY 61 GGGCTTTTAAACATCCAGAGAACAGAGAAATGTGTGACTGCAAAACAAAGGTCAGA 120
DB 319 GGTCTTCTACGGCATCTCTGAGAAATAGAAATGCGAGACTGCAAGTCAAGGGTCTCGA 378
QY 121 TGGGTAGTGTAAATTTAGTATCTTTATCTGCATGCAATGTCTGGGATTCACAGGAGT 180
DB 379 TGGCAAGTGTGAATCTAGTATCTTTATGATGATGACATGTCTGGCATTCATAGAGC 438
QY 181 CTCGGGGTACACATATCGAAGTTCGATCTGCCATCTGGACACATGCTCCCGAGCAG 240
DB 439 CTGGGGTGCACATATCTAAGTAAGATCTGCCACCTGGATACATGCTGCCAGCA 498
QY 241 GTTGCAATTATACAGTCAATGGGAAATGATAAGCAATAGTACTTGGGAACAGAGCTA 300
DB 499 GTTGCAATTATTCATCAATGGGAAACGAAAGCAATAGTATTTGGGAACAGAGCTG 558
QY 301 CCCCACAACTATGATAGTGTGMAATTTGAGAAATTTATACGTGCAAAAGTATGAAGAG 360
DB 559 CCTCTTAATGATAGTGTGMAATTTGAGAAATTTATACGTGCAAAATGATGAGGACAG 618
QY 361 AGATGGGTTCCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCCAGCAGGAAAGCGGG 420
DB 619 AGATGGGTACCAAGGAAATGGAACATCAAAATCTTCGTCCAGTTCGAGATGAGAAGAGC 678
QY 421 AAATCTGTGGAGAGAGTGGCGGGGATAGAGCATGGAATAGTAGT--AGTCTGTAA 478
DB 679 CAAGAGTCTCCGGCTAGTCTAATAGGAGTGGACATGGTTCATAGATCTTCATTTGAGCA 738
QY 479 ATTTGTTTTCAGGAGGAGGAAACATATTCAGCATCTAGACAGAAATATGTTGCTGCA 538
DB 739 AACCGTGTCTACCAAGCTCTTCGAGCAAGTTGCAATGCGAGCTTCAAGGATATCTCT 798
QY 539 CGAGATAAA-TCTTCCCTGCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAG 597
DB 799 CAGGCATCACCTCAGCCCCCGAGGTAGAAACACAGTTCGAAGGTGTTTACTCTCT 858
QY 598 AAAATGGAGTCTGCAGTACTCCAGTATGAGAGGGGAGAAACAGAGCATAAATGTTGACCA 657
DB 859 CAGTCACAGAAATCCCCCGCAAGGTTGATGCAACACCGCCTAAAGTTGAGAAGCCATCG 918
QY 658 GCATCAGATCTCCAAAGGTGGATTTTGCTATCTGATCTGTTTAAACATGCTATCAATGGAT 717

DB 919 GTTACACACACCTCTCTAAAGTTGATTATGCGCACTGATCTCTTTAAACATGTTGTCAATGGAT 978
QY 718 GATTTCGACTACAAATACCTCAGAGGCAACTCTCGGATATCTCTCGCGGATGATACTCA 777
DB 979 GGAACAAACAGAGAAAGAGTCAGCGTCA-----TCTTCAACAGCAGATTAATGGC 1026
QY 778 TGGCTGGCTTTTCACTCTGCTGGAAAGTGGTCAAAAGCGCAGAGAAATTTGTCCACAGCCAA 837
DB 1027 TGGATATGGCTTCAGTCTGCAACACCAAGTACTAGTTTCAGAGAAAGATTTCTGCCAAA 1086
QY 838 CTTCTCAGAGCAGTCTCTCCAGCTTCATCTTCTGACTTTGAGGATTTGTTTAAAGAC 897
DB 1087 CCAGCAGAAAGCAAGACCCAG-----TCTACATCTGGAATGGAAGACTTTATTTAAAGAC 1140
QY 898 ACACCTAATTTAAACACTCAACAGCACCAAAAGATGTGAAGGC-----GAT 945
DB 1141 TCACAGCTGTGCCATTATCTCAGCTCCAGCTGTTTCCCAGTTAAACGCTAAGATGAT 1200
QY 946 ATCATGAGCCTGTTTGAAGAGCAATATAGTATCGCTTTTGGCATCATCAGCAACAG 1005
DB 1201 ATCATGAGTGTGTTTGAAGAGTCCCAATATGATATCACCTTTGCTGCCCATCAACAGCAG 1260
QY 1006 GTTCTATGCTGCTCAGCAGCAAGCCCTTTTACATGCTGCGTGCAGGAAAGCTG 1057
DB 1261 CTGGCGCTCATGCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTCAAGAGCTG 1312
RESULT 8
US-10-425-115-176514
; Sequence 176514, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425, 115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 176514
; LENGTH: 2299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_92573C.1
US-10-425-115-176514
Query Match 23.4%; Score 340; DB 20; Length 2299;
Best Local Similarity 61.0%; Pred. No. 1.4e-93;
Matches 654; Conservative 0; Mismatches 385; Indels 33; Gaps 5;
QY 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCCGCATAGAAAGATTCTTGAA 60
DB 327 ATGAACGAGAAAGGTCGCTCTCCAGGAGCTCAAGCCAGCACAGAGATATTGGAA 386
QY 61 GGGCTTTTAAACATCCAGAGAACAGAGAAATGTGTGACTGCAAAACAAAGGTCAGA 120
DB 387 GGTCTTCTACGGCATCTCTGAGAAATAGAAATGCGCAGACTGCAAGTCAAGGGTCTCTGA 446
QY 121 TGGGTAGTGTAAATTTAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
DB 447 TGGCAAGTGTGAATCTAGGTATCTTTATATGATGATGATGTTCTGGCATTCATTAAGC 506
QY 181 CTCGGGGTACACATATCGAAGTTCGATCTGCACTCTGGACACATGCTGCCGAGCAG 240
DB 507 CTGGGGTGCACATATCTAAGTAAAGTCTGCCACCTTGGATACATGCTGCCAGGCA 566
QY 241 GTTGTATTTATACAGTCAATGGGAAATGATAAAGCAATAGTACTTGGGAAGCAGAGCTA 300

Db 567 GTTGCAATTATTCAATCAATGGGAACGAAAGCAAAATAGCTATTGGGAAGCAGAGCTG 626
QY 301 CCCCCAACTATGATAGAGTTGGATTTGAGATTTTATACGTGCAAGATGATGAAGAGAG 360
Db 627 CTTCTTAATCATGATAGAGTTGGATTTGAGATTTTATACGTGCAAGATGATGAAGAGAG 686
QY 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTGCAGCAGGAACGGGG 420
Db 687 AGATGGGTACCAAGGATGGACATCAAAATCTTCGTCCAGTTGTCGATGAGAGAGC 746
QY 421 AAATCTGTGGAGAGAGTGGCGGATATGAGCATAGGACATAGTAGT--AGTCTGTAA 478
Db 747 CAAGAGTCTCCGGCTAGTCTAATAGAGTGCACATGGTCTAGATCTTCAATTCAGCAA 806
QY 479 ATTTGTTTGGAGGAGGAAATATTTCCAGCATCTAGACAGAAATATATGTTGCTGCA 538
Db 807 AACCGTCTTCCAGCTCTCCGAGCAAGTTGCACATGCGCTTCAAGGATATCTCT 866
QY 539 CGAGAAATAA--TCTTCCCGTGGCTCCCAAGGACCCAGTCAGGTTATTAAGCCACAGCAG 597
Db 867 CAGGCATCACCTCAGCCCCCGAAGGTAGAAACACAGTTTCCAAAGGTGTTTACCTCT 926
QY 598 AAATGGAGTCTGAGCTACTCTCAGTAGAGAGGAGAAACAGAGCATGAATGTTGCACCA 657
Db 927 CAGTCACAGAAATCCCCCGCCAAAGTTGATGCAACACCGCCCTAAAGTTGAGAAGCCATCG 986
QY 658 GCATCAGATCTCCAAAGGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGGAT 717
Db 987 GTTACACCACTCTCTAAAGTTGATATGCGCATGATCTCTTTAAACATGTTGCAATGGAT 1046
QY 718 GATTGCTACAAATACCTCAGAGGCAACTCTCGCGATPACTCTCTCGCGATGATAACTCA 777
Db 1047 GGAAACAACAGAGAGAGTCAGCTCA-----TCTTCAACGACGATATGGC 1094
QY 778 TGGGTGGCTTTTCACTCTGCTGGAAGTGTGTAACCGGAGAGAAATTTGTACAGCCCAAG 837
Db 1095 TGGGATGGCTTCCAGTCTGCAACACAGTACCTAGTTTCAGAGAAAAAAGATTTCTGCCAAA 1154
QY 838 CTGCTGTGAGCAGTTCTCTCCAGTTCATCTCTGATCTTGTAGGATTTGTTTAAAGAC 897
Db 1155 CCAGCAGAAAGCAAGCCAG-----TCTACATCTGGATGGAAGACTTATTTAAAGAC 1208
QY 898 ACACCTAATTTAACTCAACAGCACCAGCAATGTAAGAGGC-----GAT 945
Db 1209 TCACCACTGTGCCATTTATCTCAGCTCCAGCTGTTTCCAGTAAACGCTAAGATGAT 1268
QY 946 ATCATGAGCTTTTGAAGAAGCAATATAGTATCGCTTTTGGCCATGATCAGCAACAG 1005
Db 1269 ATCATGAGTTTGTGAGAAGTCCAATATGGTATCACCTTTGCTGCCCCATCAACAGCAG 1328
QY 1006 GTTGCTATGCTCGCTCAGCAGCAAGCCCTTTACATGGCTGCGCAAGAGCTG 1057
Db 1329 CTGGCGCTCATGCTCAGCAGCAAGCTCTTTCAATGGGTGCTCTCAAGAGCTG 1380

RESULT 9

US-10-425-114-4099
; Sequence 4099, Application US/10425114
; Publication NO. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4099

; LENGTH: 2011
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700349043_FLI
US-10-425-114-4099

Query Match 23.3%; Score 338.4; DB 18; Length 2011;
Best Local Similarity 60.9%; Pred. No. 3.9e-93;
Matches 653; Conservative 0; Mismatches 386; Indels 33; Gaps 5;

QY 1 ATGAACGAGAAAGCAACGCTCTCTAAGGAGCTTAATGCCGCCATAGAAAGATTTCTTCAA 60
Db 170 ATGAACGAGAAAGCGTCCGCTCTCAAGGAGCTCAAGCCCAAGCAACAAGATATTGAA 229
QY 61 GGGCTTCTTAAACATCCAGAGAAACAGAGAAATGTCTGACTGCAAAAACAAAGGTTCCAAGA 120
Db 230 GGTCTTCTACGGCATCTCGAATAGAGAAATGCGCAGACTGCAAGTCAAGGGTCTCTCGA 289
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCAATCTTCTGGGATTCACAGGAGT 180
Db 290 TGGGCAAGTGTGAATCTAGGGATCTTTATATGCAATGACATCTTCTGGCATTCATAGAGC 349
QY 181 CTCGGGTGACACATATCGAAGGTTGATCTGCCACTCTTGGACACATGGCTCCCCGAGCAG 240
Db 350 CTGGGGGTGCACATATCTAAGGTAAAGATCTGCCACCTCGGATACATGGCTGCCAGAGCAA 409
QY 241 GTTCATTTTATACAGTCAATGGGAATGTAAAGCAATAGTCTTACTGGGAGCAGAGCTA 300
Db 410 GTTCATTTTATCAATCAATGGGAAACGAAAGCAATAGCTATTTGGGAAGCAGAGCTG 469
QY 301 CCCCCAACTATGATAGAGTTGGAAATTTTATACGTGCAAGATGATGAAGAGAG 360
Db 470 CTTCTTAATCATGATAGAGTTGGATAGAGATTTTCATCCGTGCAAAATATAGGAGACA 529
QY 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCGAGCAGAGAGCGGG 420
Db 530 AGATGGGTACCAAGGAATGGAAACATCAAAATCTTCTGTCAGTGTTCGAGATGAGAAGAGC 589
QY 421 AAATCTGTGGAGAGAGTGGCGGGATATGAGCATGAGCATAGTATAGT--AGTCTGTAA 478
Db 590 CAGAGCTCTCGGCTAGTGTATATAGAGTGGACATGTCATAGATCTTCAATTTTGGAGCAA 649
QY 479 ATTTGTTTGGAGAGAGGAAAACTATTCCAGCATCTAGAAACAAGAAAAATATGTTGTCGAA 538
Db 650 AACCGTCTTCCGCCAGCTCTTCCAGCAAGTTGTCACATGCGCTTCAAGGATATCTCTCT 709
QY 539 CGAAGATAAA--TCTTCCCGTGGCTTCCCAAGGACCCAGTCAAGTTTAAAGCCACAGCAG 597
Db 710 CAGGCATCACCTCAGCCCGGAGAGGTAGAAAACACAGTTTCCAAAGGTGTTTACCTCT 769
QY 598 AAATGGAGTCTGCAGCTACTCCAGTAGAGGGGAGAAACAAGCAGTAAATGTTGCACCA 657
Db 770 CAGTCACAGAAATCCCCCGCAAGTTGATGCAACACCGCTTAAAGTTGAGAGCCATCG 829
QY 658 GCATCAGATCTCCAAAGGTGGATTTTGTACTGATCTGTTTAACTGCTATCAATGAT 717
Db 830 GTTACACCACTCTCTAAAGTTGATTTATGCCACTGATCTCTTTAACTGTTGTCATGAT 889
QY 718 GATTGCACTACAATATCTCAGAGGCAACTCTCTGGCGATACTCTCTGCGATGATTAATCA 777
Db 890 GGAAACAACAGAGAAAGAGTCAGCTTCA-----TCTTCAACGAGCAGATAATGGC 937
QY 778 TGGGCTGGCTTTTCAAGTCTGCTGGAAAGTGGTCAACCGCAGAGAAAAATGTCACAGCCCAAG 837
Db 938 TGGGATGGCTTCCAGTCTGCAACACAGTACTAGTTTCAAGAGAAAAAGATTTCTGCCAAA 997
QY 838 CTTCTGTGAGAGCAGTTCTCTCTCCAGCTTCACTTTCTGCACTTTGAGGATTTCTTTAAGAC 897
Db 998 CCAGCAGAAAGCAGAGACCCAG-----TCTACATCTGGAATGGAAGACTTATTTAAAGAC 1051
QY 898 ACACCTAATTTAACTCAACAGCACCAGCAATGTAAGAGGC-----GAT 945

Db 1052 TCACGAGTGTGCCATTATCTCAGCTCCAGCTGTTTCCCAAGTAAACGTAAGAATGAT 1111
Qy 946 ATCATGAGCTGTTTCCAGAGCAATATAGTATCGCTTTTGGCCATGATCAGCAACAG 1005
Db 1112 ATCATGAGTGTGTTTGAAGTCCCAATATGGTATCACCCTTTCGCTGCCCATCAACAGCAG 1171
Qy 1006 GTTGTCTATCTCGCTCAGCAGCAAGCCCTTTACATGGCTGCGCAGCAAAAGCTG 1057
Db 1172 CTGGCGCTCATGTCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTCAAGCTG 1223

RESULT 10

US-10-425-114-32446
; Sequence 32446, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32446
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73346C08_FLI
US-10-425-114-32446

Query Match 23.3%; Score 338.4; DB 18; Length 2022;
Best Local Similarity 60.9%; Pred. No. 3.9e-93;
Matches 653; Conservative 0; Mismatches 386; Indels 33; Gaps 5;
Qy 1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCGCCGCATAGAAAGATTCTTGAA 60
Db 242 ATGAACGAAAGGGGTCGCTCCAGGAGCTCAACGCCAAGCAAGAAAGATATTGGAA 301
Qy 61 GGGCTTTTAAACATCCAGAGAACAGAAATGTCTGACTGTCAAAACAAAGGTCACAGA 120
Db 302 GGTCTTCTACGGCATCCTGAGAATAGAAATGCGCAGACTGCAAGTCAAAAGGGTCTCGA 361
Qy 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGGAGT 180
Db 362 TGGGCAAGTGTGAATCTAGGTATCTTTATGATGATGACATGTTCTGGCATTCATAGAGC 421
Qy 181 CTCGGGTTACACATATCGAAGTTCGATCTGCACCTCTGACACATGCTCCCGCAGCAG 240
Db 422 CTGGGGTGCACATATCTAAGTAAATCTGCCACCTGATACATGCTGCCAGCAAGCA 481
Qy 241 GTTGCAATTTATACAGTCAATAGGAAATATAAGCAAAATAGTTACTTGGGAAGCAGCTA 300
Db 482 GTTGCAATTTATCAATCAATAGGAAACGAAAGCAATAGTATTGGGAAGCAGCTG 541
Qy 301 CCCCACAACTATGATAGTGGAAATTTGAAATTTATACGTGCAAGTATGAGAGAG 360
Db 542 CTTCTTAATCAGATAGGTTGGAATAGAGATTTTCATCCGTGCAAAATATGAGGACAG 601
Qy 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCCAGCAGGAAACGGCG 420
Db 602 AGATGGTTACCAAGGAATGGAACATCAAAATCTTCTGTCAGTGTTCGATGAGAGAGC 661
Qy 421 AAATCTGTGGAGAAAGTGGCGGGAATGAGCATGGAATAGTATAGT--AGTCTGTAA 478
Db 662 CAAGAGTCTCCGCTAGTGTAAATAGAGTGGACATGGTTCATAGATCTTCAATTCAGCAA 721
Qy 479 ATTGTTTGGAGGAGGAAACATATTCACGATCTAGAACAGAAATATGTTGTCGCA 538

Db 722 AACCGTGTCTCACCAGCTCTTCCGAGCAAAAGTTGCATGCGAGCTTCAAGGATATCTCT 781
Qy 539 CGAAGATAAA-TCTTCCGTCCTCCCAAGGACCCAGTCAAGTTATAAAGCCAGCAG 597
Db 782 CAGGCATCACCTCAGTCCGCCCCGAAAGTAGAAACACAGTTCCAAAGGTGTTTCTCTCT 841
Qy 598 AAATGGAGTCTGCGAGTCTACCTCCAGTAGAGAGGAGAAACAAAGCAGTAAATGTTGCACCA 657
Db 842 CAGTCACAGAAATCCCCCGCCAAAGTTGATGCAACACCGCTTAAAGTTGAGAAGCCATCG 901
Qy 658 GCATCAGATCTCCAAAGTGGATTTTGTCTACTGATCTGTCTTAAACATGCTATCAATGAT 717
Db 902 GTTACACCACCTCTCTAAAGTTGATTAATGCACTGATCTCTTTAAACATGTTGTCAATGGAT 961
Qy 718 GATTGCACTACAAATACCTCAGAGGCAACTCTCGCGGATACCTCTCGCGGATGATAACTCA 777
Db 962 GGAACACAGAGAAAGAGTCCAGCGTCA-----TCTTCAACAGCGATTAATGGC 1009
Qy 778 TGGGCTGGCTTTTCACTCTGCTGGAAGTGGTCAACCGCAGAGAAATTTGTCAAGCCCAAG 837
Db 1010 TGGGATGGCTTCAGTCTGCAACACCACTAGTCTCAGAGAAAGAAATTTCTGCCAA 1069
Qy 838 CTGCTGAGAGCAGTTCCTCTCCAGTTCATCTTCTGACTTTGAGGATTTGTTTAAAGGAC 897
Db 1070 CCAGCAGAAAGCAAGACCCAG-----TCTACATCTGGAATGGAAGACTTATTTAAAGAC 1123
Qy 898 ACACCTAATTTAACTCAACAGCAGCAACCAAGATGTGAAGGC-----GAT 945
Db 1124 TCACGAGCTGTGCCATTATCTCAGCTCCAGCTGTTTCCCAAGTAAACGCTAAGATGAT 1183
Qy 946 ATCATGAGCTGTTTGAAGACGAATATAGTATCGCTTTTGGCCATGCTCAGCAACAG 1005
Db 1184 ATCATGAGTGTGTTGAGAGTCCAAATATGATGATACCTTTTCGCTGCCCATCAACAGCAG 1243
Qy 1006 GTTGTATGCTGCTCAGCAGCAAGCCCTTTACATGCTGCGTGCAGCGAAAGCTG 1057
Db 1244 CTGGCGCTCATGTCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTCAAGCTG 1295

RESULT 11

US-10-425-114-22038
; Sequence 22038, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22038
; LENGTH: 1709
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-022-F10_FLI
US-10-425-114-22038

Query Match 23.2%; Score 336.6; DB 18; Length 1709;
Best Local Similarity 58.9%; Pred. No. 1.3e-92;
Matches 626; Conservative 0; Mismatches 419; Indels 18; Gaps 2;
Qy 1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGAGCCGCGCATAGAAAGATTCTTGAA 60
Db 77 ATGAACGAGAAAGGCGTCCGCTCTCAGGAGGCTCAACGCCGCAAGCAAGAAATATTGGAA 136


```
QY 61 GGGCTCTTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCACAGA 120
Db 137 GGTCTCTTTCGGCTTCTGAGATAGAGAAATGTGCACTGCAAGTCANAGGGTCTCGA 196
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCAATGTTCTGGGATTCACAGGAGT 180
Db 197 TGGGCAAGTGTAAATCTGGGATCTTTATATGATGACATGTTCTGGAATTCATAGAAGC 256
QY 181 CTCGGGTACACATATCGAAGGTTGCAATCTGCACTCTGACACATGCTCCCGGAGAG 240
Db 257 CTGGGGTGCAATATCTAAGGTAAGATCTGCTACCTGGATACATGCTGGTCCGACAGAA 316
QY 241 GTTGCAATTTATACAGTCAATGGAATATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
Db 317 GTTGCAATTTATCAATCAATGGAACGAAAGCAAAATAGCAATGGAATGGAAGCAGCTG 376
QY 301 CCCCAAATCTATGATAGTGTGAATGGAATTTATAGCTGCAAGTATGATGAAGAGAG 360
Db 377 CCTCTTAATATGATAGGTTGGGATAGAGAAATTTATCCGTGCAAAATATAGGACAA 436
QY 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAACGGCGG 420
Db 437 AGATGGGTACCGAGGAATGGAACATTCGACCTTCTCCGGTGTTCGAGATGATAGAAGC 496
QY 421 AAATCTGTGGAGAGAGTGGCGGGATATGAGCATGGAATAGTAGTAGTCTCTGTAAT 480
Db 497 CAAGAGTCTCCAGCTAGTCTAATAGGATGACATCATAGATCTTCATTTGAGCAAAAC 556
QY 481 TTGTTTGGAGAGAGGAAACATATTCAGCATCTAGAACAAAGAAATATGTTGCTGCAAG 540
Db 557 CGAGCTTCAACAGCTCTCCGAGCAAGGTTGCACCTGTAGCTTCAAGGATACCTCTCAG 616
QY 541 AGAATAAATCTTCCTGGCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
Db 617 GCATCACCTCAGCCCCCAAGGTAGAACCCAGCTTCTTAAGGTGTTTCACTCTCAG 676
QY 601 ATGGAGTCTGCACTACTCCAGTAGAGAGGGAGAAACAAAGCAAGTAAATTTGCAACGCA 660
Db 677 CCACAAAATCTCTCCCAAGTTGAGGCAACACCCCTTAAGTTGAGAAGCCATCAGTT 736
QY 661 TCAGATCTCCAAAGTGTGATTTGCTACTGATCTGTTTAAATGATCAATGATGATAT 720
Db 737 GCGCGGCTCTCTAAAGTTGATATGCTACTGATCTCTTTAAATGATCAATGATGAT 796
QY 721 TCGACTACAAATACCTCAGAGGCAACTCTCGGCGATCTCTCGCGATGATAACTCATGG 780
Db 797 ACAACAGAGAAAGTCAGAGTCATCTCAA- - - - -CGATGATATGCTTGG 844
QY 781 GCTGGCTTTTCACTGCTGGAAGTGTCAACCGCAGAGAAAATTTGTCAAGCCAAAGCT 840
Db 845 GATGGCTTCCAGTCTGCACAAACAGTACCTAGCTCGGAGAAAAGATTTCTGCCAAACCA 904
QY 841 GCTGAGAGCA- - - - -GTTCTCTCAGCTTCACTCTCTGACCTTTGAGGATTTGTTTAA 894
Db 905 GCAGAAAGCATGACCCAGTCTACCTGGAATAGAAAGCTTATTTAAAGATTTGCGCAGTT 964
QY 895 GACACACCTTAATTTAACTCAACAGCACCACAAAGATGCAAGGGCGATATCATGAGC 954
Db 965 GTGTCAATATCTCAGCTCCAGCTGTTTCCCAAGTAAATGCAAGAAATATCATGAT 1024
QY 955 CTGTTTGAAGAGCAATATAGTATGCTGCTTTTGGCATGTCATCAGCAACAGGTTGCTATG 1014
Db 1025 TTGTTTGAAGTCCCAATATGCAATACCATTTGCTGCTCCACAGCAGCAGCTGGCTT 1084
QY 1015 CTCGCTCAGCAGAGCCCTTTTACATGTTGCTGCTGAGGCAAGCTG 1057
Db 1085 ATGTCTCAGCAAGCTCTTCTAATGGCTGCTCTTAAAGCTG 1127
```

RESULT 12

US-10-425-114-30303

; Sequence 30303, Application US/10425114

; Publication No. US2004003488A1

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; FILE REFERENCE: 38-21(53313)B
; NUMBER OF SEQ ID NOS: 2003-04-28
; CURRENT FILING DATE: 2003-04-28
; SEQ ID NO 30303
; LENGTH: 1802
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73008B02_FLI
US-10-425-114-30303
```

Query Match 23.2%; Score 336.6; DB 18; Length 1802;

Best Local Similarity 58.9%; Pred. No. 1.3e-92; Indels 18; Gaps 2;

Matches 626; Conservative 0; Mismatches 419;

```
QY 1 ATGAACAGAGAAAGCCAAACGCTCTCTAAGGAGCTTAATGCCGCCCATAGAAAGATTCTTGAA 60
Db 149 ATGAACAGAGAGGGCTCGTCTCCAAGGAGCTCAACGCCAAGCAACAGACATATTGCA 208
QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCACAGA 120
Db 209 GGTCTTCTCGGCTTCTGAGATAGAGAAATGTGAGATGCTCAAGTCAAGGCTCTCGA 268
QY 121 TGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCTTCTGGATTCACAGGAGT 180
Db 269 TGGCAAGTGTAAATCTTGGATCTTTATATGATGACATGTTCTGGAATTCATAGAAGC 328
QY 181 CTCGGGTACACATATCGAAAGGTTTCGATCTGCCACTCTGGACACATGGCTCCCCGAGCAG 240
Db 329 CTGGGGTGCACATATCTAAGTAAGATCTGTACCTCGGATACATGGCTGCCAGAGCAA 388
QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAATAGTTACTGGGAAGCAGAGCTA 300
Db 389 GTTGCAATTTATCAATCAATGGGAAACGAAAGCAATAGCCATTTGGGAAGCAGAGCTG 448
QY 301 CCCCAACATATGATAGAGTTGGAATTTGAGAAATTTATACGTCAAAAGTATGAAGAGAG 360
Db 449 CCTCTAATCTATAGAGTTGGATGAGAAATTTATCCGTGCAAAATATGAGGACAG 508
QY 361 AGATGGGTTTCTAGAGGGGAAAAGGCTTAGATCACCTCTAGAGTCGAGCAGGAACGGCGG 420
Db 509 AGATGGGTACCGAGGAATGGAACATTCGACCTTCTCTCGGTGTCGAGATGATAAGAGC 568
QY 421 AAATCTGTGAGAGAAAGTGGCCGGGATATGAGATGGAATAGTAGTAGTAGTCTCTGTAAT 480
Db 569 CAAGAGTCTCCAGCTAGTCTAATAGAGTGGACATCATAGATCTTCATTTGAGCAAAAC 628
QY 481 TTGTTTGAAGAGAGAAAACATTTTCCAGCATCTAGAACAAAGATAATTTGTGCTGCAACG 540
Db 629 CGAGCTTCAACAGCTCTTCGAGCAAAAGTTGCACCTGTAGCTTCAAGGATACCTCTCAG 688
QY 541 AGAATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCAAGTTATAAGCCACAGCAGAAA 600
Db 689 GCATCACTCAGCCCCCAAGGTAAGAACCCAGTCTCTAAGGTGGTTTCACTCTCTCAG 748
QY 601 ATGGAGTCTGAGCTACTCCAGTAGAGGGAGAAACAGCAGTAAATTTGTCACCGACGCA 660
Db 749 CCACAAAATCTCTCGCAAAAGTTGAGGCAACACCCCTAAAGTTGAGAACCCATCAGTT 808
QY 661 TCAGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAACTGCTATCAATGGATGAT 720
Db 809 GCGCGGCTCTCTAAAGTTGATATGCTACTGATCTCTTTAACTGTTATCAATGATGGA 868
```

Qy	721	TCGACTCAAAATACCTCAGAGGCAACTCTCTGGCGATAC	CTCTGCGCGATGATPACTCATGG	780
Db	869	ACAACAGAGAAAGAGTCAGAGTCATCTCCAAA-----	CGATGATTAATGCGCTGG	916
Qy	781	GCTGGCTTTTCAGTCTGCTGGAAAGTGTC	CAACGGCAGAGAAAATGTGCACAGCCAAGCCT	840
Db	917	GATGGCTTCCAGTCTGCACAACCACTACCTAGCTCGG	GGGAAAGAATCTGCCAACAACCA	976
Qy	841	GCTGAGAGCA-----GTTCTCTCCAGCTTCATCTTCT	GTGACTTTGAGGATTTGTTTAAAG	894
Db	977	GCAGAAAGCATGACCCAGTCTACACCTGGAAATAGAAG	CTTATTTTAAAGATTTGCCCGATT	1036
Qy	895	GACACACTTAATTTAAACAACCTCAACAAGCAC	CAAAAGATGTGAAAGCGCATATCATGAGC	954
Db	1037	GTGTCAITATCCTCAGCTCCAGCTGTTTCCCAAGTAA	ATGTCAAAGAATGATATCATGAGT	1096
Qy	955	CTGTTTGAGAAAGCAATATAGTATCGCTTTTCCCAT	TGCATCAGCAACAGGTTGCTATG	1014
Db	1097	TTGTTTGAGAAAGTCATATATGGCATCACCATTCG	CTGTCCAACAGCAGCAGCTGGCGTTT	1156
Qy	1015	CTCGCTCAGCAGCAAGCGCCCTTTACATGGCTG	CAGCGAAAGCTG	1057
Db	1157	ATGTCCTCAGCAGCAAGCTCTCTTAATGGCTGCTCT	TTAAAGCTG	1199

RESULT 13

```

US-10-425-115-176512
; Sequence 176512, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 176512
; LENGTH: 1928
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577 92571C.1
; US-10-425-115-176512

```

Query Match	23.2%;	Score 336.6;	DB 20;	Length 1928;
Best Local Similarity	58.9%;	Prod. No. 1.4e-92;		
Matches 626;	Conservative 0;	Mismatches 419;	Indels 18;	Gaps 2;
Qy	1	ATGNAACAGAAAGCCAAAGCTCTCTTAAGGAGCTTAATGATGCCCCCATAGAAAGATCTTGGAA	60	
Db	275	ATGAACAGAGAGGGCTCGGTCTCCAGAGGAGCTCAAGCCCAAGCACAGAGATATTGGAA	334	
Qy	61	GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCACAAAACAAAAGGCTCCAAGA	120	
Db	335	GGTCTTCTTCGGCTTCTCTGAGAAATAGAGAAATGTCAGACTGCAAGTCAAAGGGTCTCTCGA	394	
Qy	121	TGGGCTAGTGTTAATTAGGTATCTTTATCTGCATATGCAATGTTCTGGGATTCACAGAGAT	180	
Db	395	TGGGCAAGTGTTAATCTTGGGATCTTTATATATGATGACATGTTCTGGAAATTCATAGAAGC	454	
Qy	181	CTGGGGGTACACATATCGAAGGTTGATCTTCGCCACTCTGGACACATGGCTCCCGGAGCAG	240	
Db	455	CTGGGGGTGCACATATCTAAGGTAAAGTCTGCTACCCCTGGGATACATGGCTGCCAGAGCAA	514	
Qy	241	GTTTCATTTTATACAGTCAATCGGGAATTCATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA	300	
Db	515	GTTTCATTTTATTCGAATCAATCGGGAACGAAAGCAAAATAGCCATTGGGAAGCAGAGCTG	574	
Qy	301	CCCCCAAACTATGATAGAGTTTGGAAATTTGAGAAATTTTATACGTGCAAGATGATGAAGAGAAAG	360	

RESIT.T 14

```

RESULT 14
US-10-425-114-17368
Sequence 17368, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 17368
LENGTH: 1915
TYPE: DNA
ORGANISM: Zea mays

```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3069-017-C3_FLI
US-10-425-114-17368

Query Match      23.1%; Score 335; DB 18; Length 1915;
Best Local Similarity 58.8%; Pred. No. 4.3e-92;
Matches 625; Conservative 0; Mismatches 420; Indels 18; Gaps 2;

QY 1 ATGACGAGAGAACCAAGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
DB |||||
QY 277 ATGACGAGAGAGCGCTCGCTCTCAAGAGCTCAAGCCAGCAAGAGATATTGGAA 336
DB |||||
QY 61 GGGCTTTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGGTCCAAGA 120
DB |||||
QY 337 GGTCTTCTTCGGCTTCTGAGATAGAGATGTGCAGACTGCAAGTCAAAGGTCTCTCGA 396
DB |||||
QY 121 TGGGCTAGTGTTAATTTAGTATCTTTATCTGATGCAATGTCTGGGATTCACAGGAT 180
DB |||||
QY 397 TGGGCAAGTGTTAATCTCGGATCTTTATATGTATGACATGTTCTTGGAAATTCATAAGC 456
DB |||||
QY 181 CTCGGGTTACACATATCGAAGGTTCGATCTGCCACTCTGCACATGGCTCCCGAGCAG 240
DB |||||
QY 457 CTGGGGTGCACATATCTAAGGTAAAGTCTGTACCCTGGATACATGGCTGCCAGAGAA 516
DB |||||
QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
DB |||||
QY 517 GTTGCAATTTATCCAAATCAATGGGAAACGAAAGCAAAATAGCCATTGGGAAGCAGCTG 576
DB |||||
QY 301 CCCCCAACTATGATAGAGTTGGAATTTATAGTGTGTCGCAAAAGTATGAAGAGAG 360
DB |||||
QY 577 CTTCTTAACCTATGATAGGTTGGGATAGAGAAATTTTCATCGTGCAAAATATGAGGACAG 636
DB |||||
QY 361 AGATGGTCTTCAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCAGCAGAGAAAGCGGG 420
DB |||||
QY 637 AGATGGGTACCGAGGAATGGAACATTCGGACCTTCTCCGGTTCGAGATGAGAGAGC 696
DB |||||
QY 421 AAATCTGTGGAGAGAGTGGCGGGATATGAGCATGAGCATAGTAGTAGTCTCTTAAT 480
DB |||||
QY 697 CAAGAGTCTCCAGTAGTCTAATAGGATGGAGATCATAGATCTTCAATTCAGCAAAAC 756
DB |||||
QY 481 TTGTTTGAGAGAGGAAACTATTCAGCATCTAGAAACAGAAATATGTTGCTGCAAGC 540
DB |||||
QY 757 CGTGCTTCCAGCTCTTCCGAGCAAGTTGCACATGTAGTCTCAAGGATACCTCTCAG 816
DB |||||
QY 541 AGAATAAATCTTCCGCTGCTCCCAAGGACCCAGTCAGGTTAATAAGCCACAGCAGAA 600
DB |||||
QY 817 GCATCACCTCAGCCCCCAAGGTAGAACCCACAGTTCTTAAGGTGTTCACTCTCAG 876
DB |||||
QY 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGGAGAAACAAAGCAGTAAATGTTGCCACAGCA 660
DB |||||
QY 877 CCACAAAATCTCTGCCAAAGTTGAGGCAACACCCCTTAAAGTTGAGAAGCCATCAGTT 936
DB |||||
QY 661 TCAGATCTCCAAAGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
DB |||||
QY 937 GCGCCGCTCTCTAAAGTTGATATGCTACTGATCTCTTTAAACATGTTATCAATGATGGA 996
DB |||||
QY 721 TCGACTACAAATACCTCAGAGGCACTCTCTGGCGATCTCTCCGCGATCAATCACTG 780
DB |||||
QY 997 ACACAGAGAGAGTTCAGATCTCTCAAA-----CGATGATATGCTGG 1044
DB |||||
QY 781 GCTGGCTTTTCAGTCTGCTGGAAGTGTCAACCGCAGAGAAATTTGTCAACGCCAGCT 840
DB |||||
QY 1045 GATGGCTCCAGTCTGCACAAACAGTACCTAGCTCGGAGGAAAGATTTCTGCCAAACCA 1104
DB |||||
QY 841 GCTGAGAGAGATTTCTCTCAGCTTCACTCTCTGA-----CTTGAGGATTTGTTAAG 894
DB |||||
QY 1105 GCAGAAAGCATGACCCAGTCTACATCTGGAATAGAGACTTATTTAAAGATTTCCGAGTT 1164
DB |||||
QY 895 GACACACCTTAATTTAAACACTCAACAGCAGCAAAAGATGTCAAGGGCGATATCATGAGC 954
DB |||||
QY 1165 GTGTCTATATCTCAGCTCAGCTGTTTCCCAAGTAAATGCAAGAAATGATATCATGAT 1224
DB |||||
QY 955 CTGTTTGAGAGAGCAATATAGTATCGCCTTTTGGCCATGATCATCAGCAACAGGTGCTATG 1014
DB |||||
```

```
Db 1225 TTGTTTGAGAAAGTCCAAACATGCGCATCGCAATTCGCTGTCCAAACAGCAGCGCTGGCGTTT 1284
QY 1015 CTCGCTCAGCAGCAGCCCTTTTACATGCTGCTCAGCGAAAGCTG 1057
Db 1285 ATGCTCAGCAGCAAGCTCTTCTTAATGCTGCTCTTAAAGCTG 1327

RESULT 15
US-10-425-115-176513
; Sequence 176513, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 176513
; LENGTH: 2221
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2221)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_92572C.1
US-10-425-115-176513

Query Match      23.0%; Score 333.4; DB 20; Length 2221;
Best Local Similarity 58.7%; Pred. No. 1.5e-91;
Matches 624; Conservative 0; Mismatches 421; Indels 18; Gaps 2;

QY 1 ATGACGAGAGAACCAAGCTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
DB |||||
QY 261 ATGACGAGAGAGCGCTCGCTCTCCAGAGGCTCAAGCCAGGACACAGAGATATTGGAA 320
DB |||||
QY 61 GGGCTTTTAAACATCCAGAGAACAGAGAAATGTGCTGATGCAAAACAAAAGGTCCAAGA 120
DB |||||
QY 321 GGTCTTCTACGGCATCTCTGAGATAGAGAAATGCGCAGACTGCAAGTCAAAGGTCTCTCGA 380
DB |||||
QY 121 TGGGCTAGTGTTAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAT 180
DB |||||
QY 381 TGGGCAAGTGTGAATCTAGGTATCTTTATATGCAATGATGTTCTGGCATTCATAGAGC 440
DB |||||
QY 181 CTCGGGTTACACATATCGAAGTTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCAG 240
DB |||||
QY 441 CTGGGGTGCACATATCTAAGGTAAAGTCTGCTACCTGGATACATGCTGCCAGCA 500
DB |||||
QY 241 GTTGCAATTTATACAGTCAATCGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
DB |||||
QY 501 GTTGCAATTTATCCAATCAATCGGAAACGAAAGCAAAATAGCCATTGGGAAGCAGAGCTG 560
DB |||||
QY 301 CCCCCAACTATGATAGAGTTGGAAATTTGAGAAATTTTATACGTGCAAGTATGAAGAGAG 360
DB |||||
QY 561 CTTCTAATCTATGATAGGTTGGATAGAGAAATTTATCTCCGTGCAAAATATAGGAGCAAG 620
DB |||||
QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGATCGAGCAGGAAACGGGG 420
DB |||||
QY 621 AGATGGGTACCGAAGATGGAACATTCGACCTTCTCCGTGTTTCGAGATGAGAGAGC 680
DB |||||
QY 421 AAATCTGTGGAGAGAGTGGCGGGGATATGAGCATGGAATAGTAGTAGTCTCTTAAT 480
DB |||||
QY 681 CAAGAGTCTCCAGCTAGTGTAAATAGGAGTGGACATCATAGATCTTCTGTTTGGAGCAAAAC 740
DB |||||
QY 481 TTGTTTGAGAGAGGAAACCTATTTCAGCATCTAGAACAGAAATAATGTTGCTGCAAGC 540
DB |||||
```

```
Db 741 CGTGCTTCCAGAGCTCTCCGAGCAAGTTGCACATGTAGTCTCAAGGATACCCCTCTCAG 800
Qy 541 AGAATAAATCTTCCGTGCTCTCCAGAGCAAGCTCAGGTTATTAAGCCACAGCAGAA 600
Db 801 GCATCACCTTCAGGCCCCAAAGGTAGAACCAACAGGTTCTTAAGGTGGTTTCACTCTCTCAG 860
Qy 601 ATGGAGTCTGCAGCTACTCTCAGTAGAGGGAGAAACAAAGCAGTAATAATGTTGCACAGCA 660
Db 861 CCACAAAATCTTTCGCAAGTTGAGGCAACACCCCTAAAGTTGAGAGCCATCAGTT 920
Qy 661 TCAGATCTCCAAAGGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
Db 921 GCGCGCTCTCTAAAGTTGATTTGATGCTACTGATCTCTTTAACAATGTTATCAATGATGGA 980
Qy 721 TCGAGTACAAATACCTCAGAGGCAACTCTCTGGGATACTCTGCGGATGATAACTCATGG 780
Db 981 ACAAAGAGAAAGAGTCAGAGTCATCTCAAA-----CGATGATAATGCCTGG 1028
Qy 781 GCTGGCTTTTCAGTCTGCTGGAAGTGTCAAACGGCAGAGAAATTTGCACAGCCAAAGCCT 840
Db 1029 GATGCTTCCAGTCTGCACAACCAAGTACCTAGCTCGGAGGAAAGATTCGCCAAACCA 1088
Qy 841 GCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGA-----CTTTGAGGATTTGTTTAA 894
Db 1089 GCAGAAAGCATGACCCAGTCTACATCTGGATAGAAAGACTTATTTAAAGATTCGCCAGTT 1148
Qy 895 GACACACCTTAATTTAACTCAACAGCACCAGAAAGATGTGAAGGCGATATCATGAGC 954
Db 1149 GTGTCAATATCTCAGTCTCAGCTGTTTCCCAAGTAAATGCAAGAAATGATATCATGAGT 1208
Qy 955 CTGTTTGAGAGCAATATAGTATCGCTTTTGCCATGCATCAGCAACAGGTTGCTATG 1014
Db 1209 TTGTTTGAGAGTCCAAATATGCAATCACCATTGCTGTCCACACAGCAGCTGCGGTTT 1268
Qy 1015 CTCGCTCAGCAGCAAGCCCTTTTACATGGCTGAGGCAAGCTG 1057
Db 1269 ATGTCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTTAAAGCTG 1311
```

Search completed: September 27, 2005, 07:23:19
Job time : 3616 secs